2014

Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus

Michael Worobey
University of Arizona, worobey@email.arizona.edu

Guan-Zhu Han
University of Arizona

Andrew Rambaut
University of Edinburgh

Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus

Michael Worobey[a,1], Guan-Zhu Han[a], and Andrew Rambaut[b,c,d]

[a]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721; [b]Institute of Evolutionary Biology and [c]Centre for Infection, Immunity, and Evolution, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom; and [d]Fogarty International Center, National Institutes of Health, Bethesda, MD 20892

Edited by Neil M. Ferguson, Imperial College London, London, United Kingdom, and accepted by the Editorial Board April 8, 2014 (received for review December 30, 2013)

The source, timing, and geographical origin of the 1918–1920 pandemic influenza A virus have remained tenaciously obscure for nearly a century, as have the reasons for its unusual severity among young adults. Here, we reconstruct the origins of the pandemic virus and the classic swine influenza and (postpandemic) seasonal H1N1 lineages using a host-specific molecular clock approach that is demonstrably more accurate than previous methods. Our results suggest that the 1918 pandemic virus originated shortly before 1918 when a human H1 virus, which we infer emerged before ∼1907, acquired avian N1 neuraminidase and internal protein genes. We find that the resulting pandemic virus jumped directly to swine but was likely displaced in humans by ∼1922 by a reassortant with an antigenically distinct H1 HA. Hence, although the swine lineage was a direct descendent of the pandemic virus, the post-1918 seasonal H1N1 lineage evidently was not, at least for HA. These findings help resolve several seemingly disparate observations from 20th century influenza epidemiology, seroarcheology, and immunology. The phylogenetic results, combined with these other lines of evidence, suggest that the high mortality in 1918 among adults aged ∼20 to ∼40 y may have been due primarily to their childhood exposure to a doubly heterosubtypic putative H3N8 virus, which we estimate circulated from ∼1889–1900. All other age groups (except immunologically naive infants) were likely partially protected by childhood exposure to N1 and/or H1-related antigens. Similar processes may underlie age-specific mortality differences between seasonal H1N1 vs. H3N2 and human H5N1 vs. H7N9 infections.

phylogeny | cohort immunity | pathogenicity | virulence | reassortment

The influenza pandemic of 1918–1920 killed an estimated 50 million people, most during a single wave late in 1918 (1, 2). Its origin, epidemiology, and pathogenesis are still puzzling (3, 4). Unusually for influenza A virus (IAV), which typically kills primarily infants and the elderly, young adults aged about 20–40 y suffered extensive mortality, which peaked in 25- to 29-y-olds (1, 2). The same virus was comparatively mild in those only slightly older or younger (Fig. 1). The very elderly, moreover, suffered less influenza-related mortality during the pandemic than in 1911–1917 (5). Notably, the virus was clinically unremarkable in >95% of patients (2), and almost all fatalities were caused by secondary bacterial pneumonia (6, 7). Any explanation must therefore account not just for the mortality peak in 20- to 40-y-olds but also for the mortality "troughs" in the very elderly and in children ∼5–15 y of age (Fig. 1), as well as for the typical postinfluenza complications (rather than acute viral pathogenesis) that killed most victims. The rapid reversion to more usual IAV mortality patterns by the early 1920s must also be explained.

Current hypotheses of the origin of the 1918 H1N1 virus range from an introduction of all eight genome segments from an avian source shortly before 1918 (8) to reassortment involving progenitor viruses supposedly circulating in humans and swine for decades before 1918 (9). Here, by using a molecular clock model that explicitly allows different evolutionary rates in different hosts, which we recently demonstrated to be essential for accurate inference of the timing and directionality of IAV host jumps (10), we reconstruct the evolutionary origins of the 1918 pandemic H1N1 virus, the classic swine H1N1 influenza virus, and the postpandemic seasonal H1N1 lineage. We then evaluate numerous observations from epidemiology, seroarcheology, and molecular evolution in light of these findings. Our results suggest that the childhood exposure of various age cohorts to different HA and neuraminidase (NA) subtypes was a key factor underlying the age-specific patterns of fatality not only in 1918 but also during other pandemics and seasonal influenza epidemics.

Results

Evidence That H1 Emerged in Humans Before ∼1907, Not in 1918. Host-specific local clock (HSLC) analyses of H1 *HA* data reveal two crucial points (Fig. 2 and *SI Appendix*, Fig. S1*A*). First, human H1 emerged from an avian source considerably earlier than 1918, sometime after the human + avian H1 most recent common ancestor (MRCA) at 1901 [95% credible interval (CI): 1895–1907] but before the MRCA of the pandemic and seasonal H1 lineages at 1907 (1903–1910). Second, the classic swine influenza lineage is nested within the 1918 genetic diversity of human H1, whereas the seasonal human H1 *HA* is distantly related to the pandemic *HA*. This pattern indicates that the swine influenza lineage emerged directly from the human pandemic virus but that postpandemic seasonal H1N1 did not. Rather, there is strong phylogenetic evidence that it descended from a distinct H1 lineage that shared a common ancestor with the 1918 pandemic *HA* in ∼1907 (i.e., both the seasonal and pandemic HA genes were evidently drawn from >10 y of accumulated

Significance

The origin of the 1918 pandemic influenza A virus (IAV) and the reasons for its unusual severity are two of the foremost biomedical mysteries of the past century. We infer that the virus arose via reassortment between a preexisting human H1 IAV lineage and an avian virus. Phylogenetic, seroarcheological, and epidemiological evidence indicates those born earlier or later than ∼1880–1900 would have had some protection against the 1918 H1N1 virus, whereas many young adults born from ∼1880–1900 may have lacked such protection because of childhood exposure to an antigenically distinct H3N8 virus. Our findings suggest that better understanding of how initial exposure shapes lifetime immunity may enhance the prediction and control of future IAV pandemics and seasonal epidemics.

Author contributions: M.W. designed research; M.W., G.-Z.H., and A.R. performed research; M.W., G.-Z.H., and A.R. analyzed data; and M.W. devised the age-specific mortality model and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. N.M.F. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

See Commentary on page 7892.

[1]To whom correspondence should be addressed. E-mail: worobey@email.arizona.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1324197111/-/DCSupplemental.

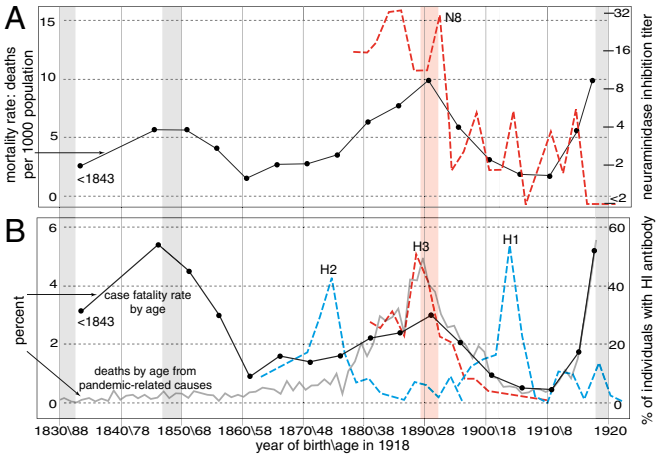


Fig. 1. Mortality in 1918 and seroarcheological patterns. (*A*) Influenza and pneumonia mortality (solid line) in different age groups in the United States in 1918 (data from ref. 1), and N8 neuraminidase inhibition titers (legend at right) in sera from different age groups (data from ref. 11). Vertical bars indicate the pandemics of 1830–1833, 1847–1850, 1889–1893, and 1918–1920. (Ages/birth years are indicated below *B*). (*B*) Case-fatality ratios in 1918 (in black) (data from ref. 1) and percentages of deaths by age from pandemic-related causes in Ontario, Canada (in gray) (data from ref. 33). Also shown are percentages of sera from different age groups with HA inhibition antibody titers (*SI Appendix*).

H1 *HA* genetic diversity circulating in the human population in 1918). This extensive diversity is consistent with the 6–12 y of circulating seasonal H1 diversity seen in the 1930s, 1940s, and 1950s, but it is distinctly deeper than the just 1–2 y of circulating *HA* diversity sampled in all later pandemic years [Fig. 2 (green rectangles) and *SI Appendix*, Fig. S1*B*].

Additional analyses indicate that these inferences are robust to the exclusion of potentially laboratory-adapted human and swine strains from the 1930s and several short 1918 *HA* sequence fragments (7) (*SI Appendix*, Fig. S2) and to the inclusion of human H1N1 sequences only from 1918 to 1957 (before the reemergence of H1N1 in 1977) (*SI Appendix*, Fig. S3). The results cannot be explained as an artifact arising from an adaptive burst of amino acid-changing substitutions in the rapidly evolving HA globular head domain because an independent analysis of the conserved stalk domain yields similar results (*SI Appendix*, Fig. S4) and there

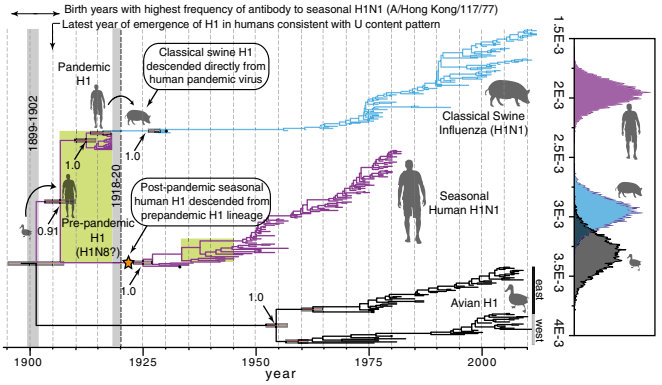


Fig. 2. Maximum clade credibility (MCC) tree of the H1 subtype of HA. (*Right*) Clade-specific rate distributions (in substitutions per site per year). (*Left*) Time window of the pandemic of 1918–1920 is shown with a gray bar, as is that of the putative pandemic around 1900. The posterior probability of each node and the 95% CIs on node dates are shown. The orange star indicates the H1 variant that evidently gave rise to postpandemic seasonal H1N1 lineage in humans. The widths of the green rectangles indicate the comparable extent of H1 genetic diversity in 1918 and 1945 (more than 10 y in each case).

is no evidence of episodic diversifying selection (*SI Appendix*, Fig. S5). Accordingly, both third codon position sites (at which 97% of substitutions are synonymous) and silent third position sites give virtually identical timing and topology estimates as the full dataset (*SI Appendix*, Fig. S6).

These results are consistent with an avian-to-human movement of H1 in the first decade of the 20th century. This decade is an interval during which a putative influenza pandemic occurred, in ~1900 (11–13). It was retrospectively identified on the basis of increased pneumonia and influenza mortality in North America, England, Ireland, and elsewhere but is not universally considered a bona fide pandemic. It was thought to have occasioned the emergence of an H3 pandemic virus until subsequent reinterpretation of H3 seroarcheological observations compellingly indicated H3's emergence during the 1889–1893 "Russian" influenza pandemic (14). Analysis of uracil (U) content, which tends to increase after avian IAV segments are transmitted to mammals (15), corroborates these phylogenetic results. Human N1 U content in the 1918 sequence is within the avian range (*SI Appendix*, Fig. S7), and there are avian sequences with a virtually identical base composition to the A/Brevig Mission/1/1918 *NA* at all four nucleotides (e.g., A/duck/NZL/76/1984). Thus, seven of the eight segments of the 1918 human virus, encoding the polymerase proteins (PB2, PB1, PA); the nucleocapsid protein (NP); the matrix proteins (M1/2); and the nonstructural proteins (NS1/2) (10) as well as NA, exhibit avian-like U content. The HA-encoding segment is the only one in the genome of the 1918 virus with a U content significantly above the avian distribution ($P = 0.0003$), consistent with it emerging in humans earlier than the other segments (*SI Appendix*, Fig. S7). We estimate this host jump would have had to occur by 1905 or earlier to allow sufficient time to reach such a high U content by 1918.

HA Seroarcheology Corresponds to Phylogeny. Many findings from seroarcheology and epidemiology, although overlooked in recent years, also point to this earlier period, and not to 1918, for the introduction of H1 into the human population (11, 13, 16–24) (Fig. 1 and *SI Appendix*, *SI Text*): (*i*) Seroarcheological results point to an introduction of H1 between 1896 and 1907, not in 1918 (16–18, 21, 25) (note the peak in antibodies to seasonal H1 in those born around 1904 in Fig. 1*B*, a pattern that seems inexplicable if H1 were new to humans in 1918); (*ii*) seroarcheology suggests the disappearance of H3 not in 1918 but shortly after ~1900 (11, 13, 18, 19) (Fig. 1*B*); and (*iii*) mortality patterns in 1968–1970 indicate that the childhood exposure of those born after ~1900 was to a non-H3 virus, which offered no protection from the 1968 H3N2 virus (23, 24), whereas those born before 1900 enjoyed considerable protection from prior exposure to H3 antigens.

The prominent peaks in antibody titers against H3 and N8 in those born in and around 1889 (Fig. 1) and the low mortality of those aged >70 y in 1968–1970 provide compelling evidence that the 1889–1893 pandemic was caused by an H3N8 virus (11, 13, 18, 26). It is widely assumed that the 1889 virus circulated until the 1918 pandemic. However, only about half of those born in 1893 were primed with H3 (14). Given the high attack rates of IAV (22), this observation seems incompatible with an H3 virus circulating until 1918 because it would require that half of 25-y-olds remained immunologically naive to IAV in 1918. We therefore hypothesize that the period around 1900, not 1918, occasioned the disappearance of H3 and the reemergence of H1 in humans. However, such seroepidemiological results can be difficult to interpret with high precision, not least because particular viruses can greatly influence the conclusions of such studies. For example, it is conceivable (although we think unlikely) that an H3N8 virus circulated from 1889 right up until ~1918 but that its HA underwent such extensive antigenic drift after ~1900 that antibodies against it fail to bind to either 1968 human H3 or 1963 equine H3 HA antigens, which are phylogenetically very divergent (10).

Interestingly, Kendal et al. (11) found that although H3 antibodies were rare or absent in those born after 1900, N8 "persisted in a virus prevalent from the early 1900s until 1916 or 1917" (also Fig. 1*A*). Based on the intersection of the phylogenetic and

seroarcheological evidence, we therefore propose that this N8 NA may have been carried over from a putative H3N8 virus that circulated from 1889 until ~1900 to an H1N8 virus that emerged in the first decade of the 20th century. If so, prior IAV exposure of the ~1900–1918 cohort would have been to this putative H1N8 virus. The seroarcheological data (Fig. 1*B*) suggest that H1 had largely displaced H3 by about 1905. This idea is consistent with the phylogenetic evidence (Fig. 2); the high U content of the 1918 HA (*SI Appendix*, Fig. S7); the lack of protection of those born after ~1900 to the 1968 H3 virus (23, 24); and, as described below, the mortality patterns in 1918.

Origin and Emergence of the 1918 Virus. The time of the most recent common ancestor (TMRCA) of H1 in humans, at 1907 (1904–1910) (Fig. 2), significantly predates the TMRCA of the human and avian N1 lineages, at 1913 (1911–1916) ($P = 0.008$; *SI Appendix*, Fig. S8). Hence, and this is a crucial point, there is strong phylogenetic evidence that N1 was introduced from an avian virus only after H1 was already established in humans. The human H1 lineage also significantly predates the human + avian TMRCA of *PB1* at 1914 (1911–1916) ($P = 0.006$; *SI Appendix*, Fig. S9), suggesting that PB1, too, was transmitted from an avian host via reassortment with a human prepandemic H1 lineage. The remaining internal genes also evidently arose from a Western hemispheric lineage of avian influenza virus (10). Their nearly identical phylogenetic patterns and their avian-like U contents (10) suggest they may have been transmitted during the same reassortment event as *PB1*. The most parsimonious scenario is that a single event in ~1915 (see below) brought together a human H1 with the remaining seven avian segments, possibly via an H7N1 virus.

If the avian segments had a single source, the human-avian virus reassortment event must have occurred after ~1914 but before the TMRCA of the human and swine lineages of each segment. To estimate these dates, we conducted an analysis of each segment, including only human and swine H1N1 viruses, allowing a separate rate for each host using the HSLC model. *SI Appendix*, Fig. S10 shows that the root node TMRCAs for the swine + human trees are remarkably consistent across segments: PB2, 1915 (1913–1917); PB1, 1915 (1913–1918); PA, 1914 (1912–1916); HA, 1914 (1912–1916); NP, 1914 (1910–1916); NA, 1914 (1911–1916); M1/2, 1914 (1909–1917); and NS1/2, 1915 (1911–1918) (full trees are shown in *SI Appendix*, Fig. S11). This result suggests an emergence of the 1918 pandemic virus ancestor in ~1915, with a window from 1913 to 1916 overlapping among all eight 95% CIs and each segment. Thus, with the important exception of the preexisting human H1 HA, our conclusions support an avian origin of the virus shortly before 1918. This conclusion suggests that its genesis was similar to that of the pandemic viruses of 1957 and 1968, which emerged via reassortment of avian and preexisting human viruses within human hosts (26, 27), but with seven segments of avian origin acquired compared with three in 1957 and two in 1968.

Classic Swine Influenza and Postpandemic H1N1. For every segment, the human + swine root node date (or, in the case of HA, the TMRCA of the swine and pandemic H1 lineage) significantly postdates the TMRCA of the pandemic and seasonal human H1 lineages (*SI Appendix*, Fig. S10). A separate analysis of HA excluding numerous 189-nt sequence fragments (7) similarly reveals a strongly supported swine + 1918 (human) clade (posterior probability = 1.0) dated at 1917 (1916–1918) (*SI Appendix*, Fig. S2). In other words, the human H1 HA genetic diversity is significantly older than the MRCA of the human and swine lineages for all eight segments (*SI Appendix*, Fig. S10). These observations suggest that swine H1N1 descended from the human virus, not vice versa. There is no evidence supporting the hypothesis (9) that reassortment between decades-old human and swine viruses played a role in the origin of the 1918 pandemic lineage. In addition, unlike the conclusions of Smith et al. (9), these phylogenetic results agree with on-the-ground observations in 1918 that the disease was new to swine and was caused by the same agent as the concurrent human influenza pandemic (16, 28).

Unlike *HA*, the remaining segments of the seasonal H1N1 virus were evidently direct descendants of the pandemic virus (*SI Appendix*, Fig. S10). This finding suggests that during the pandemic, reassortment occurred between the pandemic lineage and a cocirculating, antigenically distinct H1 virus, creating the seasonal H1N1 ancestor. This scenario parallels the emergence of the 1946–1947 H1N1 lineage and the associated worldwide vaccine failure (29). In both cases, rather than evolving directly from its predecessor via antigenic drift, a new lineage with distinct antigenic properties evidently arose via the acquisition of an "old" and rare homosubtypic H1 HA variant by intrasubtype reassortment. The replacement of the pandemic HA by a heterologous H1 HA offers a simple resolution to the long-standing conundrum that antibodies to the swine/pandemic virus HA suddenly disappeared in those born after ~1922 (16, 17, 20, 30, 31), even though seasonal H1N1 obviously continued to circulate (*SI Appendix*, *SI Text*). However, less parsimonious scenarios, such as a separate cross-species introduction of the seasonal H1 HA from birds or even from other mammalian hosts such as horses, in or shortly after 1918, cannot be formally excluded (*SI Appendix*, *SI Text* and Fig. S12). Recovery of archival IAV genomes from ~1907–1917 and the 1920s might definitively resolve these questions.

Model to Explain the 1918 Mortality Patterns. Elderly individuals may have been protected from the 1918 virus by childhood exposure to an H1N1-like virus (5). We estimate that H1 and the H2 + H5 lineage diverged from a common ancestor near the time of the 1830 pandemic (*SI Appendix*, *SI Text* and Figs. S13 and S14). Moreover, protection was clearly greatest in those born before 1834 (5) (Fig. 3*A*), implicating the 1830–1833 pandemic virus, which would have primed the majority of that age group. If an H1-like virus emerged in 1830, it would likely have been positioned near one of the orange stars close to the root of the tree in *SI Appendix*, Fig. S13. Those primed as children between 1830 and 1889 by this HA lineage would likely have had considerable protection against the 1918 HA, comparable to that exhibited during the 2009 H1N1 pandemic by those born before 1957 (32), based on the similar genetic distances separating the childhood and pandemic virus HA in each case (*SI Appendix*, Figs. S13 and S14).

It is likely that the majority of those born worldwide shortly before and during 1889–1893 were infected by the 1889 pandemic H3N8 virus (12). A progressively larger percentage of those born either earlier or later, however, would have experienced initial childhood exposure to the putative H1 viruses that immediately preceded (1830–1889) or followed (~1900–1918) the 1889 virus. Based on observations of the 1893 cohort, about half of whom were primed with H3 (14), we assume in Fig. 3*B* that ~0% of newborns, ~50% of 7-y-olds, and ~100% of 14-y-olds would already have been exposed to prepandemic seasonal strains in 1847, 1889, 1900, and 1918. Although these numbers are undoubtedly inexact, we believe they are reasonable for developing a qualitative model as described below.

We propose that putative 1830–1847 H1N1, 1847–1889 H1N8, 1889–1900 H3N8, and 1900–1918 H1N8 IAV lineages determined the childhood exposure patterns of the various age groups alive in 1918 (Fig. 3*B*, *Lower*). In Fig. 3*B* (*Upper*), we depict the protection expected against the 1918 H1N1 virus for each birth-year cohort from 1830 to 1918 under this scenario. All else being equal, the 1830 cohort (who were 88 y old in 1918) would have had the best protection against the 1918 H1N1 virus due to childhood exposure to a doubly homosubtypic virus. The 1847 cohort (71-y-olds) would have had intermediate protection due to the presence of H1 HA antigens in their childhood virus, but a heterosubtypic NA. The 1889 cohort (29-y-olds) would have had the least protective immunity from childhood, almost all having been primed with a heterosubtypic H3 HA and

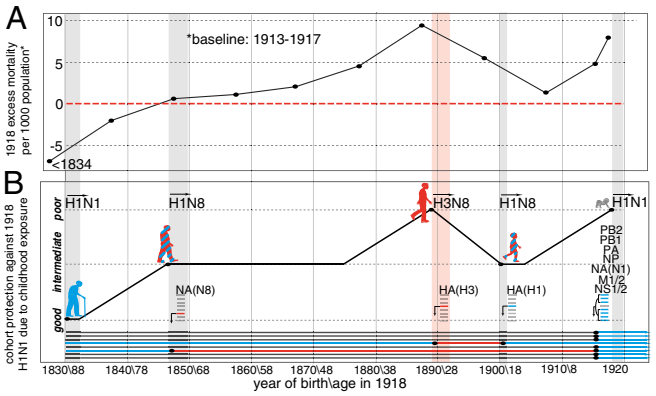


Fig. 3. Excess mortality in 1918 and the childhood exposure/cohort immunity model. (*A*) Age group-specific annual excess mortality due to pneumonia and influenza in 1918 (data from ref. 5). (Year of birth and age in 1918 are indicated below *B*.) (*B*) Expected cohort protection due to childhood exposure. Each segment of the 1918 H1N1 genome is shown in blue. Putative H1- and N1-like genes in the 1830, 1847, and 1900 genomes are also shown in blue, whereas putative heterosubtypic genes (H3 and N8) in the 1847, 1889, and 1900 viruses are shown in red. The human silhouettes are colored by putative childhood exposure to HA and NA antigens matched or mismatched to the 1918 H1N1 (+/+, blue; +/–, blue/red; –/–, red), whereas the newborn is colored gray, indicating no prior exposure. The pandemics of 1847, 1889 (red), 1900, and 1918 are indicated with vertical bars.

a heterosubtypic N8 NA. The 1900 cohort (18-y-olds) would have had intermediate protection, having been exposed to a homosubtypic HA and heterosubtypic NA. Finally, most infants, with the proportion increasing to ∼100% with declining age, would have lacked protection because they would have had no prior exposure whatsoever to IAV antigens.

The expected protection of the cohorts born in the years between the pandemics would follow the curve shown in Fig. 3*B*, given the assumptions detailed above. Half of those aged 7 y in 1889, for example, would already have been primed by the 1847–1889 virus before the 1889 pandemic; the remainder would have been primed by the 1889–1900 virus, leaving the 1882 birth year cohort (36-y-olds) with protection in 1918 midway between the 1847–1875 and 1889 birth-year cohorts (Fig. 3*B*).

The excess mortality patterns in 1918 (Fig. 3*A*) closely match the expected protection curve based on this cohort immunity model (Fig. 3*B*). Indeed, the actual peak in mortality among young adults occurred precisely in those born from 1889 to 1893 (1) (25- to 29-y-olds; Fig. 1), the only age group in 1918 whose childhood exposure would have been almost exclusively to a fully heterosubtypic virus (H3N8) (also ref. 33). Note that the peak in excess mortality in 1918 (Fig. 3*A*) closely overlaps the peaks in both N8 and H3 antibody responses in birth years near 1889 (Fig. 1 *A* and *B*). In particular, the percentage of deaths due to pandemic influenza-related causes in various age groups in several Canadian cities, from the most thorough analysis to date of mortality data from 1918 (33), shows a striking congruence with the presence of H3 antibody titers (Fig. 1*B*), one that is difficult to dismiss as random coincidence.

Importantly, no single event, including the 1889 pandemic, can explain the overall pattern. We propose that it was the aggregate exposure of the various cohorts to different pandemic viruses, and crucially also to the seasonal influenza lineages in interpandemic periods, that set the stage for what unfolded in 1918. This scenario provides a straightforward potential explanation for the inflection points in the mortality-by-age curve in 1918, including the observation that the most elderly cohort suffered lower mortality in 1918 than in 1911–1917. The switch to H1N1 may have resulted in a virus to which their childhood exposure provided better protection than it did to the putative H1N8 strain immediately preceding the pandemic. The same switch would have had the opposite effect in young adults exposed in childhood to H3N8, limiting the usefulness of their N8 NA-directed antibody responses and leaving them with comparatively ineffective antibody protection against both major antigenic glycoproteins of the pandemic H1N1 virus. The age groups on either side of the "H3N8" cohort would have enjoyed considerable protection against H1N1 due to their childhood exposure to homosubtypic HA antigens (Fig. 3*B*), whereas the youngest individuals lacked any prior exposure and associated protection in 1918.

It is also possible that an absence of prior immunity to the new avian M2 and NP proteins of the 1918 virus contributed to the severity of the 1918 pandemic. M2 immunity is associated with decreased viral replication in the lungs and less severe disease (34), and preexisting immunity to NP decreases susceptibility to secondary bacterial pneumococcal infection (35). The M2 and NP proteins of the 1918 virus may have been considerably divergent from the previous human variant, which likely predated the homogenization of IAV internal genes after the selective sweep in these genes that occurred in the late 1800s (10).

We can conceive of two mechanisms whereby the childhood exposure of different age groups could have shaped the mortality patterns in 1918. First, a mechanism akin to original antigenic sin (OAS) (36) may have interfered with immune responses in some of those infected in 1918 (33, 37), peaking in those exposed to the 1889 virus. Although OAS has been traditionally considered a within-subtype phenomenon (36, 38–40), it is plausible that interactions between heterosubtypic viruses could also occur (41). Indeed, Masurel (42) reported that when immunized with an H3N2 vaccine, about 5% of individuals primed in childhood by H1N1 yielded strong HA inhibition antibody responses to H1N1, without any appearance of antibody responses to H3N2 virus. We also speculate that exposure to H1 HA stalk antigens could have resulted in unprotective (OAS-mediated) recall of antibodies to H3 HA stalk epitopes in some H3N8-primed individuals. Such misdirected immune responses could have had dire consequences in 1918 for those initially infected by H3N8. Even if most or all 20- to 40-y-olds in 1918 had already been exposed to the putative H1 virus circulating between ∼1900 and 1918, we speculate that their initial exposure to an H3 virus might nevertheless have interfered with their immune responses to the 1918 HA. This process, combined with weak or nonexistent immunity to the newly emerging NA, M2, and NP proteins in the 1918 virus, may have rendered this subgroup at particularly high risk for severe disease and more vulnerable than young adults in any pandemic since 1918. These other pandemics have occurred against a backdrop of widespread prior exposure of the human population to a homosubtypic NA (1968, 2009) and/or HA (1977, 2009), or to a heterosubtypic but phylogenetically closely related HA (*SI Appendix*, Fig. S13) with a very similar HA stalk (1957) (33).

The alternative is that an unlucky subset of individuals had been exposed only to heterosubtypic H3 HA antigens before 1918 (i.e., they had escaped infection by the putative H1N8 virus that circulated from ∼1900–1918). Unlike virtually all older and younger age cohorts (except infants), such an H1-naive subset of 20- to 40-y-olds would have been unable to benefit from anamnestic immune responses to H1 antigens. Importantly, immunopathology is not invoked in this scenario: Heterosubtypic prior infection would provide positive but limited protection, better than being completely immunologically naive but worse than prior exposure to N1 and/or H1 antigens. There is ample evidence from animal models that this effect might be expected: Prior exposure to a heterosubtypic influenza A virus (but not to influenza B virus) is associated with lower death rates compared with completely naive hosts (35, 43). However, it is also associated with higher viral lung titers and enhanced pneumonia and death rates compared with prior exposure to a homosubtypic virus. The observation that even higher mortality than in 20- to 40-y-olds was observed in 1918 in groups that were largely immunologically naive, young infants (Fig. 1*B*) and isolated populations (e.g., a staggering 22% mortality in Samoa) (44), suggests that exposure to H3N8, although hardly

optimal, may nevertheless have been considerably better than no prior IAV exposure at all.

It is important to note that unlike antigenic imprinting this "H1-naive" hypothesis requires that a sizeable proportion of 20- to 40-y-olds remained unexposed to H1 before 1918, although most younger individuals (e.g., the excess mortality trough in those ∼15 y of age in 1918) were exposed. Although this idea might be unrealistic, we note that if an H1 virus indeed emerged in the early 20th century, it was mild enough to go unnoticed at the time (11–13); in addition, although seasonal influenza virus attack rates peak at the age of 2–3 y and reach an average of 20.3% per year among children under 5 y of age, they are much lower for adults of working age: just 6.6% on average, including influenza B virus infections (45). Thus, it is plausible that a quarter or more of the 20- to 40-y-olds alive in 1918 could have remained unexposed even if an H1N8 virus had been circulating seasonally for 10–15 y before 1918.

Discussion

We hypothesize that childhood exposure to an H3N8 virus may have made some young adults in 1918 a sort of temporal counterpart to highly vulnerable geographically isolated populations, inducing suboptimal immunity that tilted the odds in favor of secondary infection with the wide range of bacterial pathogens that cause most influenza-related mortality. This small wedge of the population may have had ineffective immunity not only to some antigens that are currently targets for "universal" vaccines (M2 and NP, newly emerged from avian influenza in 1918) but, uniquely, also against both major antigenic glycoproteins and the conserved HA stalk domain. Antibodies targeting the HA stalk provide powerful protection against severe disease, but there is little or no cross-protection between phylogenetically divergent group 1 HA subtypes (e.g., H1 in 1918) and group 2 subtypes (e.g., H3 in 1889) (46). If this model is correct, then current medical interventions, especially antibiotics and vaccines against several pneumonia-causing bacteria, could be expected to reduce mortality dramatically if we were faced today with an otherwise similar set of pandemic ingredients.

If childhood exposure of different age groups is indeed a key predictor of outcome to a pandemic strain (i.e., if antigenic imprinting and not just the intrinsic virulence of the virus shapes mortality patterns), then current approaches to studying influenza pathogenesis might need to be rethought. For instance, using immunologically naive animals to characterize the pathogenicity of the reconstructed 1918 virus may be methodologically problematic, because the pathogenesis of the actual pandemic virus appears to have been profoundly affected by the prior immunity experienced by different age cohorts, a potentially crucial factor not reflected in such an experimental design. This idea could also help explain why genomic analyses of large datasets, such as for the 2009 pandemic H1N1 virus, have not yet identified "virulence factors" that can explain why the virus causes a life-threatening infection in some people but is asymptomatic in others. In short, it may be useful to frame questions about pandemic IAV pathogenicity in terms of how well (or poorly) prior exposure protects different age groups from the >20% general mortality that can occur in "virgin soil" populations (44, 47).

Finally, our findings suggest that childhood exposure of different age groups to distinct influenza virus variants may also strongly influence age-related mortality patterns during seasonal epidemics, as well as to H5N1 and H7N9 viruses. We hypothesize that the current severity of seasonal H3N2, which was remarkably mild among the elderly cohort of 1968 (22) but now kills ∼18-fold more patients aged >65 y than H1N1 (48, 49), may be, in part, a consequence of antigenic imprinting rather than the intrinsic virulence of the virus. The current elderly cohort (unlike the elderly cohort in 1968, who had been primed in ∼1889–1900 by an H3 virus) was primed exclusively by H1N1 (*SI Appendix*, *SI Text*). We predict that as the 1968 H3N2-primed cohort begins to replace the H1N1- and H2N2-primed cohorts among those >65 y of age, H3N2-dominated epidemics may diminish in frequency and severity and H1N1-dominated epidemics may increase (assuming these subtypes are still cocirculating in future decades).

Prior exposure of some age cohorts to H7N9 and H5N1 clearly cannot explain the age-specific mortality patterns seen with these viruses, because they are not thought to have circulated previously in humans; all age groups have had prior exposure only to heterosubtypic HA. However, childhood exposure to group 1 (H1, H2) vs. group 2 (H3) HA antigens, ones that are either shared or not shared with the group 1 HA of H5N1 or the group 2 HA of H7N9, is remarkably predictive of disease severity: We find that virtually all fatalities from H5N1 (group 1) have occurred in younger patients initially exposed to H3N2 (group 2) and, conversely, that almost all H7N9 (group 2) mortality has occurred in older individuals initially infected by H1N1 or H2N2 (group 1). The opposing patterns of age-specific mortality/protection with H5N1 and H7N9 are highly statistically significant and are almost perfect mirror images on either side of the group 1-to-group 2 HA transition that occurred with the emergence of H3N2 in 1968 (*SI Appendix*, *SI Text* and Fig. S15). We hypothesize that anamnestic recall of immune responses to the HA stalk antigens of initial childhood exposure may underlie this pattern by predisposing patients to severe disease when they encounter H5N1 or H7N9 viruses with group-mismatched HA but strongly protecting them when these viruses have group-matched HA proteins.

Immunization strategies that mimic the apparently powerful lifetime protection afforded by initial childhood exposure might dramatically reduce mortality due to both seasonal and novel IAV strains. Better understanding of cohort immunity effects, which may be more pervasive and powerful than previously appreciated, might thus lead to improved understanding of IAV pathogenesis and to better prediction, prevention, and control of both seasonal and pandemic influenza.

Materials and Methods

IAV Sequence Data Preparation. We collected all IAV full-length sequences from humans, birds, and pigs encoding the H1, H2, and H5 subtypes of HA and the N1 subtype of NA. Identical sequences and apparent recombinants and other problematic sequences were excluded. For each gene, a subset of sequences of a size amenable to molecular clock analyses (∼300 sequences) was sampled, preserving the most basal sequences in the major clades and reducing the number of overrepresented recent sequences so that sampling across different years was fairly even. Because the effective sampling time of post-1977 to pre-2009 human H1N1 is 27 y earlier than the actual sampling date, we shifted the dates accordingly (10). The full-length swine and human IAV sequences from the alignments of the PB2, PB1, PA, NP, M1/2, and N1/2 genes from a study by Worobey et al. (10) were used for the analyses summarized in *SI Appendix*, Figs. S9 and S10.

Phylogenetic Analyses. We analyzed these IAV alignments with the HSLC model as described (10) using a Gaussian Markov random field Bayesian skyride coalescent tree prior and a general time reversible + gamma substitution model. Each major host group was allowed its own rate in the HSLC model. For the analysis of the H1, H2, and H5 subtypes, all of the avian sequences were assumed to evolve at the same rate or were allowed independent rates, with similar results in each case (*SI Appendix*, Fig. S14); the human H2 and human H1 clades were allowed their own rates. We ran analyses for 50 million steps in most cases and used Tracer v1.5 to ensure effective sample size values >200. We used TreeAnnotator to infer and annotate MCC trees. To test the robustness of the deep, pre-1918 divergence time of the human H1 lineage, as well as the clustering of the 1918 sequences with the classic swine influenza lineage rather than with the postpandemic seasonal human H1N1 lineage, we conducted several additional analyses of H1 datasets. These analyses included (*i*) exclusion of the 189-nt HA fragments from 1918, as well as laboratory strains of IAV from both humans and swine from the 1930s; (*ii*) subsampling at most one sequence per host lineage per year; (*iii*) subsampling only sequences sampled before the extinction of H1N1 in 1957; (*iv*) separate analysis of the HA stalk domain (sites 1–150 and sites 921–1,698); and (*v*) analysis including only the 565 third-position sites and the subset of 503 silent third-position sites. [We used MacClade v4.08a (50) to visualize all amino acids substitutions along the MCC tree and then determined which were due to substitutions at the third codon position by referring to the genetic code and the nucleotide

alignment. One hundred thirteen of 4,107 third-position substitutions along the MCC tree were nonsynonymous.]

U Content Analyses. We compared the U content of the 1918 *HA* and *NA* sequences with the range observed in avian viruses (*SI Appendix*, Fig. S7). We estimated an upper bound on when the 1918 *HA* sequence emerged in a mammalian host using the approach described by Worobey et al. (10), calculating how long a sequence starting at the average U content among avian strains would take to increase to the U content value observed in the 1918 sequence, assuming the rate of U content increase in human H3 *HA* (because it appears that the H1 lineage was approaching an asymptote between 1918 and 1957). The overall H3 substitution rate (10) is slightly higher than that of H1 (Fig. 2), so this assumption likely provides a conservative estimate of the upper bound (i.e., if the rate of U content increase in H1 were slightly lower than in H3, this discrepancy would suggest the entry into humans was slightly earlier than our estimate of ~1905). The upper and lower range estimates were determined using the upper and lower 95% confidence interval values for the avian U content distribution. A *P* value for a test of the hypothesis that the avian-to-human jump predated 1918, based on U content, was calculated as the proportion, out of 10,000 replicates, in which the year drawn from the above-mentioned distribution was greater than (i.e., postdated) 1918.

Tests for Adaptive Evolution. We used the random effects branch-site model (51) for detecting episodic diversifying selection (EDS) in the H1 *HA* phylogeny. We included representative sequences from each host lineage to permit a search for evidence of EDS on the branch between each host, and within each host after putative host jumps (*SI Appendix*, Fig. S5).

Tests of Whether Within-Human H1 *HA* Diversity Predates Between-Host Diversity in Other Genes. A *P* value for a test of the hypothesis that the within-human diversity of the H1 subtype of *HA* predates the human + swine + avian N1 *NA* diversity was calculated by drawing a date from the human H1 TMRCA posterior density and a date from the multihost N1 TMRCA posterior density, and then determining the proportion of 10,000 replicates for which the N1 date was earlier than the H1 date. The same approach was used for tests of whether the within-human H1 diversity predates the human + swine diversity within N1 and each of the internal genes (*SI Appendix*, Fig. S10) and for a test of whether the within-human H1 diversity predates the human + swine + avian *PB1* diversity (*SI Appendix*, Fig. S9).

ACKNOWLEDGMENTS. We thank two anonymous reviewers for insightful criticisms and comments. This work was supported by a grant from the David and Lucile Packard Foundation (to M.W.) and Grant 092807 from the Wellcome Trust (to A.R.). The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement 278433-PREDEMICS and European Research Council Grant Agreement 260864.

1. Frost WH (1919) The epidemiology of influenza. *JAMA* 73(5):313–318.
2. Taubenberger JK, Morens DM (2006) 1918 Influenza: The mother of all pandemics. *Emerg Infect Dis* 12(1):15–22.
3. Morens DM, Taubenberger JK (2012) 1918 influenza, a puzzle with missing pieces. *Emerg Infect Dis* 18(2):332–335.
4. Morens DM, Fauci AS (2007) The 1918 influenza pandemic: Insights for the 21st century. *J Infect Dis* 195(7):1018–1028.
5. Luk J, Gross P, Thompson WW (2001) Observations on mortality during the 1918 influenza pandemic. *Clin Infect Dis* 33(8):1375–1378.
6. Morens DM, Taubenberger JK, Fauci AS (2008) Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: Implications for pandemic influenza preparedness. *J Infect Dis* 198(7):962–970.
7. Sheng ZM, et al. (2011) Autopsy series of 68 cases dying before and during the 1918 influenza pandemic peak. *Proc Natl Acad Sci USA* 108(39):16416–16421.
8. Taubenberger JK, et al. (2006) Molecular virology: Was the 1918 pandemic caused by a bird flu? Was the 1918 flu avian in origin? (Reply). *Nature* 440(7088):E9–E10.
9. Smith GJ, et al. (2009) Dating the emergence of pandemic influenza viruses. *Proc Natl Acad Sci USA* 106(28):11709–11712.
10. Worobey M, Han GZ, Rambaut A (2014) A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* 508(7495):254–257.
11. Kendal AP, Minuse E, Maassab HF, Hennessy AV, Davenport FM (1973) Influenza neuraminidase antibody patterns of man. *Am J Epidemiol* 98(2):96–103.
12. Patterson KD (1986) *Pandemic Influenza, 1700-1900: A Study in Historical Epidemiology* (Rowan & Littlefield, Totowa, NJ).
13. Davenport FM, Minuse E, Hennessy AV, Francis T, Jr. (1969) Interpretations of influenza antibody patterns of man. *Bull World Health Organ* 41(3):453–460.
14. Dowdle WR (1999) Influenza A virus recycling revisited. *Bull World Health Organ* 77(10):820–828.
15. Rabadan R, Levine AJ, Robins H (2006) Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. *J Virol* 80(23):11887–11891.
16. Shope RE (1936) The incidence of neutralizing antibodies for swine influenza virus in the sera of human beings of different ages. *J Exp Med* 63(5):669–684.
17. Francis T, Magill TP (1936) The incidence of neutralizing antibodies for human influenza virus in the serum of human individuals of different ages. *J Exp Med* 63(5):655–668.
18. Masurel N, Mulder J (1966) Studies on the content of antibodies for equine influenza viruses in human sera. *Bull World Health Organ* 34(6):885–893.
19. Masurel N (1969) Relation between Hong Kong virus and former human A2 isolates and the A-EQU12 virus in human sera collected before 1957. *Lancet* 1(7601):907–910.
20. Masurel N, Heijtink RA (1983) Recycling of H1N1 influenza A virus in man—A haemagglutinin antibody study. *J Hyg (Lond)* 90(3):397–402.
21. Rekart M, Rupnik K, Cesario TC, Tilles JG (1982) Prevalence of hemagglutination inhibition antibody to current strains of the H3N2 and H1N1 subtypes of influenza A virus in sera collected from the elderly in 1976. *Am J Epidemiol* 115(4):587–597.
22. Bodewes R, et al. (2011) Prevalence of antibodies against seasonal influenza A and B viruses in children in Netherlands. *Clin Vaccine Immunol* 18(3):469–476.
23. Simonsen L, Reichert TA, Miller C (2004) The virtues of antigenic sin: Consequences of pandemic recycling on influenza-associated mortality. *Int Congr Ser* 1263:791–794.
24. Schoenbaum SC, Coleman MT, Dowdle WR, Mostow SR (1976) Epidemiology of influenza in the elderly: Evidence of virus recycling. *Am J Epidemiol* 103(2):166–173.
25. Mulder J, Masurel N (1958) Pre-epidemic antibody against 1957 strain of Asiatic influenza in serum of older people living in the Netherlands. *Lancet* 1(7025):810–814.
26. Webster RG, Laver WG (1972) The origin of pandemic influenza. *Bull World Health Organ* 47(4):449–452.
27. Kawaoka Y, Krauss S, Webster RG (1989) Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J Virol* 63(11):4603–4608.
28. Koen JS (1919) A practical method for the field diagnoses of swine diseases. *Am J Vet Med* 14(9):468–470.
29. Nelson MI, et al. (2008) Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. *PLoS Pathog* 4(2):e1000012.
30. Andrewes CH, Laidlaw PP, Smith W (1935) Influenza: Observations on the recovery of virus from man and on the antibody content of human sera. *Br J Exp Pathol* 16(6):566–582.
31. Davenport FM, Hennessy AV, Drescher J, Mulder J, Francis T, Jr. (1964) Further observations on the relevance of serologic recapitulations of human infection with influenza viruses. *J Exp Med* 120(6):1087–1097.
32. Adalja AA, Henderson DA (2010) Original antigenic sin and pandemic (H1N1) 2009. *Emerg Infect Dis* 16(6):1028–1029.
33. Gagnon A, et al. (2013) Age-specific mortality during the 1918 influenza pandemic: Unravelling the mystery of high young adult mortality. *PLoS ONE* 8(8):e69586.
34. Fan J, et al. (2004) Preclinical study of influenza virus A M2 peptide conjugate vaccines in mice, ferrets, and rhesus monkeys. *Vaccine* 22(23-24):2993–3003.
35. Haynes L, et al. (2012) Immunity to the conserved influenza nucleoprotein reduces susceptibility to secondary bacterial infections. *J Immunol* 189(10):4921–4929.
36. Francis T (1960) On the doctrine of original antigenic sin. *Proc Am Philos Soc* 104(6):572–578.
37. Shanks GD, Brundage JF (2012) Pathogenic responses among young adults during the 1918 influenza pandemic. *Emerg Infect Dis* 18(2):201–207.
38. Davenport FM, Hennessy AV (1956) A serologic recapitulation of past experiences with influenza A; antibody response to monovalent vaccine. *J Exp Med* 104(1):85–97.
39. Webster RG, Kasel JA, Couch RB, Laver WG (1976) Influenza virus subunit vaccines. II. Immunogenicity and original antigenic sin in humans. *J Infect Dis* 134(1):48–58.
40. Kim JH, Skountzou I, Compans R, Jacob J (2009) Original antigenic sin responses to influenza viruses. *J Immunol* 183(5):3294–3301.
41. Waffarn EE, Baumgarth N (2011) Protective B cell responses to flu—No fluke! *J Immunol* 186(7):3823–3829.
42. Masurel N (1976) Swine influenza virus and the recycling of influenza-A viruses in man. *Lancet* 2(7979):244–247.
43. Schulman JL, Kilbourne ED (1965) Induction of partial specific heterotypic immunity in mice by a single infection with influenza A virus. *J Bacteriol* 89(1):170–174.
44. Shanks GD, Hussell T, Brundage JF (2012) Epidemiological isolation causing variable mortality in Island populations during the 1918-1920 influenza pandemic. *Influenza Other Respi Viruses* 6(6):417–423.
45. Molinari N-AM, et al. (2007) The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine* 25(27):5086–5096.
46. Steel J, et al. (2010) Influenza virus vaccine based on the conserved hemagglutinin stalk domain. *mBio* 1(1):10.1128/mBio.00018-10.
47. Wolfe RJ (1982) Alaska's great sickness, 1900: An epidemic of measles and influenza in a virgin soil population. *Proc Am Philos Soc* 126(2):91–121.
48. Thompson WW, et al. (2004) Influenza-associated hospitalizations in the United States. *JAMA* 292(11):1333–1340.
49. Thompson WW, et al. (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA* 289(2):179–186.
50. Maddison DR, Maddison WP (2005) MacClade 4: Analysis of Phylogeny and Character Evolution, Version 4.08a. Available at http://macclade.org. Accessed April 28, 2011.
51. Kosakovsky Pond SL, et al. (2011) A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* 28(11):3033–3043.

SI Appendix

Supplementary Text

Details of seroarcheologic results depicted in Fig. 1B. The H1 results are from data from ref 1 (HI titres ≥100 against A/Hong Kong/117/77). The H2 results from data from ref 2. The H3 results come from combining data from Fig. 2 of ref 3 (HI titres of 9-15 against an equine H3N8 virus, "A/EQUI 2/Richelieu/63", in human sera collected in 1958) with data from Fig. 1 of ref 4 (HI titres ≥20 against a human H3N2 virus, "A2/Hong Kong/68", in human sera collected in 1958), with results pooled by age group and combined into a single curve.

Seroarcheology corroborates pre-1918 emergence of H1 and extinction of H3. The highest antibody titres in an age group reflect the dominant antigens of the virus responsible for their initial childhood exposure (*5, 6*). In sera collected in 1935, both Shope (*7*) and Francis and Magill (*8*) found that protection against H1 antigens peaked not in those born just before 1918 but in those born between 1896 and 1905 (Fig. 1 in ref 7; Chart 2 in ref 8). Rekart *et al.* (*9*) and Masurel and Heijtink (*1*) found the identical pattern forty years later: in sera collected in 1976 and 1977, respectively, HI antibody assays with H1N1 viruses revealed peak H1 protection in those born between 1896 and 1905 (*9*) and 1896 and 1907 (*1*) (Fig. 2 in ref 9; Fig. 1 in ref 1).

Specifically, Masurel and Heijtink (*1*) found in 728 sera collected from patients with birth years between 1883 and 1930 that the highest percentage with HI antibodies to seasonal H1N1 viruses was in those born in 1903-4 (their results with A/Hong Kong/117/77 are superimposed on the 1918 case fatality curve in Fig. 1B: "H1"). If H1 had actually been first introduced to the human population in the 1918 pandemic one would expect to find the greatest percentage of H1 antibodies in sera collected from those born in or just prior to 1918. However, ~90% those born in 1903-4 had HI titres ≥50 against A/Hong Kong/117/77, while just ~10% of those born closer to 1918 (birth year 1909-10) did (*1*).

Furthermore, there is no strong seroarcheological evidence that H3 circulated for long after ~1900 (*2-4, 10-13*). For example, in 913 sera collected in 1956-57, Masurel (*2*) found that only those born before 1897 exhibited high HI antibody titres either to human H3 or equine H3 HA antigens. And in sera collected in 1963, while 50 of 435 taken from those born prior to 1903 showed antibodies against equine H3 antigens, only 2 of 434 from those born later did (*3*).

There is also no evidence from mortality patterns in 1968-1969 that H3 circulated after 1900: unlike the strong protection in those born just prior to, or during, the 1889-93 pandemic, those younger than 70 (born after 1896) experienced no discernable protection from excess mortality upon exposure to the 1968 H3N2 virus (*11*). Attack rates in 1968-69 were also about three times higher for those born after 1899 than those born prior to 1890 (*12*). These observations are not consistent with continued circulation of an H3 virus up until the 1918 pandemic.

Dowdle (*5*) noted, based on seroarcheological findings with H1 (*14*) and H3 (*2*), that about half of those born in 1893 had been primed with the 1889 H3, and half with H1, which he assumed had emerged in 1918, 25 years after 1893. We contend that it is untenable to argue that a virus with an average annual attack rate of 20-30% in children (*15*) would have left 50% of 25 year-olds immunologically naïve to an H3 virus that supposedly circulated up until the 1918 pandemic: this would imply an unrealistic annual attack rate averaging just 2% over a quarter century. It seems doubtful that there could have been more than a few percent of 25-year-olds in 1918 with no prior exposure whatsoever to influenza A virus (except in the most remote locations). Dowdle's observation, on the other hand, is easily reconciled with the hypothesis that H3 was replaced by H1 around 1900, and that ~50% of those born in 1893 were not exposed to H3 antigens prior to the extinction of that virus and were primed with H1 after it emerged in ~1900.

Classical swine influenza and postpandemic H1N1. To our knowledge the earliest evidence of swine influenza in 1918 comes from Dr. Grant Munger, an inspector from the Cedar Rapids, Iowa Division of Hog Cholera Control of the Bureau of Animal Industry who observed herds of swine ill with influenza in western Illinois, in August (note 1 in ref 7).

The evidence that the classical swine virus in 1918 was a direct descendant of the pandemic virus, but the postpandemic seasonal virus was not (at least in HA), makes sense of some confusing observations from

seroarcheology—including the abrupt disappearance of antibodies to the swine virus HA in those born after about 1922 (*1, 7, 8, 16, 17*), even though this same cohort showed a high frequency of potent antibodies to WS/33 and other seasonal H1N1 viruses. As Andrewes *et al.* stated in 1935 (*16*): "The age distribution of antibodies to the porcine virus is now seen to be of some interest; the absence of neutralizing sera in children under ten may be due to complete or partial dying out of this particular strain of virus as a cause of human disease during the last decade." Our phylogenetic results (Fig. 2) are in precise agreement with these early observations: the obvious inference is that children born after 1922 or so indeed were not exposed to the pandemic virus HA, which had by then been displaced by an antigenically and phylogenetically distinct seasonal H1N1 variant (orange star in Fig. 2), one that the seroarcheological results suggest may have been introduced near 1900.

Crucially, the precipitous decline in antibodies to H1 HA among those born from ~1918-1922, which has been put forward as a seroarcheological model for identifying pandemics (*5*), is observed only when antigens of classical swine influenza virus, rather than seasonal human H1N1, are used in HI assays (*1, 7, 16*). Our results imply that the steep drop in antibodies to the swine virus H1 actually reflects the process of a formerly widespread HA variant (the pandemic virus) rapidly going extinct as it was displaced by the reassortant seasonal H1N1 virus, with its related, but distinct, H1 HA. We believe our phylogenetic results provide a logical resolution to the sudden disappearance of antibodies to the pandemic/swine virus HA in the 1920s (*7, 16*) (even though H1N1 continued to circulate until 1957), as well as to the seroarcheological evidence that the HA of the seasonal H1N1 that circulated after ~1922 emerged near 1900, rather than 1918 (*1*).

The immunological backdrop of pandemic mortality patterns in 1918. There is seroarcheologic evidence of antibody responses against H2 antigens among the cohort born in the 1870s (*2*) (also see Fig. 1B). In light of our phylogenetic results, we suspect these responses likely reflect cross-reactivity to a putative 1830-1889 H1-like HA rather than to H2 per se, which our results show did not even diverge from its common ancestor with H5 until ~1875 (Figs. S13 and S14). Masurel (*14*) showed that >70% of those born between 1867-1886 had HI antibody titres ≥100 to H1 HA antigens of A/Swine/15/1930 (Fig. 1 in ref 14). This is a higher percentage than in the slightly younger age group exposed in childhood to the 1889 virus (*14*), suggesting the pre-1889 age groups were indeed exposed to H1-like antigens prior to the emergence of the 1889 virus (see Fig. 1B). Shope (*7*) found the same pattern in the 1930s: sera from 100% of those born before 1875 neutralized A/Swine/15/1930 (note the bimodal distribution in Fig. 1 of ref 7, with peak percentages of protection against H1 in those born in 1896-1905 and those born prior to 1875, i.e. either side of the 1889 pandemic).

There was also much higher mortality among >50-year-olds in 1889-90 than there was in 1918, strongly suggesting that those born before ~1840 experienced childhood exposure to a non-H3, non-N8 virus. Altogether, it seems highly plausible that an H1N1-like virus was associated with both 1830 and 1918, and that an H1-like HA circulated from 1830-1889, leaving only a short window from ~1889-1900 when an H1-like virus did not circulate in humans (Fig. S13). Much the same thing happened after 1918, with H1N1 circulating until 1957, re-emerging in 1977 and 2009, and H1N1 being absent only between 1958-1976. As shown in Fig. 1, the disappearance of an H1 HA left the ~1889 cohort with a sharp peak in heterosubtypic (with respect to H1N1) anti-HA and NA antibodies (H3 and N8, respectively), while slightly older and younger groups exhibited marked H1 or H2 HA protection.

We speculate that the 1830 H1N1-like virus acquired N8 by reassortment in or shortly before 1847, leading to that pandemic, which coincided with an extensive epizootic in horses all across Europe (*18*). This putative H1N8 variant may have remained as the seasonal strain until the 1889 pandemic, which would account for the evidence of antibodies to H1/H2 in those born between 1847 and 1889. Those primed by such a virus, with a homosubtypic HA and heterosubtypic NA, would be expected to have suffered intermediate excess mortality in 1918 (higher than the cohort primed by the putative 1830 H1N1-like virus, but lower than the cohort whose initial exposure was to the 1889 H3N8 virus).

The 1889-93 pandemic, the first in modern times, had a global scope not seen in prior influenza pandemics, and it was second only to 1918 in severity over the last two centuries. Revolutionary developments in transportation, especially modern steamships that led to an enormous increase in the volume and speed of intercontinental maritime travel, and railway networks that connected port cities and inland populations, allowed influenza to move between and across continents at unprecedented speeds (*19*). In light of this, as well as the high

incidence of disease measured during several successive waves (*19*), it is likely that most individuals born shortly before and during 1889-93 were primed by the 1889 pandemic virus.

Fan *et al.* (*20*) showed in a mouse model that M2 immunity reduces viral replication in the lungs during the entire course of infection. If so, then absence of natural immunity to M2 would likely increase viral replication in the lungs. This may be a contributing factor for the general severity of the 1918 pandemic. Our phylogenetic analyses make it clear that a new M1/2-encoding segment was acquired just prior to the pandemic from an avian virus. The prior M2 may have been so divergent that cross-protection with the 1918 M2 was poor, leading to a more lethal virus until herd immunity to the new M2 was achieved. Recovery of pre-1915 virus sequence could permit a test of this hypothesis, but the protection of the oldest age groups in 1918 suggests to us that lack of immunity to M2 or other minor antigens (like NP) may not, on its own, have been a decisive factor in 1918.

Cohort-specific immunity in 1918 and other pandemic years. The idea that young adults suffered unusually severe outcomes in 1918-1920 because of childhood exposure to H3 and N8 antigens fits with the observation that they died primarily of typical post-influenza complications. Any individual with a poor immune response to a current influenza infection is at higher risk of severe or fatal disease. Indeed, the logic behind this argument is no different than that underpinning the enterprise of protecting individuals from severe and fatal influenza-related disease by eliciting protective immune responses with vaccines. Since prior exposure to H3 protected the 1889-1900 cohort against H3N2 in 1968-70 (*11*), it stands to reason that the same age group would have been especially vulnerable to the antigenically mismatched virus in 1918, compared to other age groups exposed in childhood to N1 and/or H1 antigens. This is similar in kind if not degree to individuals <70 in 1968, <20 in 1977, and <52 in 2009 (*11, 15*). These younger age groups, primed in childhood with heterosubtypic HA or NA antigens, all suffered more excess mortality than older age groups who were apparently protected by childhood exposure to viruses with partial or complete overlap in HA and NA subtypes.

Were clinical attack rates in 1918 of about 28% overall and 30% in adults 20-40 years of age (*21*) consistent with the possibility of a pre-pandemic interval of H1 HA circulation? This is lower than observed in some interpandemic (seasonal) epidemic years when virtually all adults have had prior exposure. Prior exposure to H1 HA would only be expected to provide sterilizing immunity (and thus produce very low attack rates) against homologous HA variants, and even a year or two of antigenic drift is often sufficient to create heterologous HA. (That is why vaccines against seasonal influenza A viruses are usually reformulated every year or two). Our phylogenetic results suggest that the circulating pre-pandemic H1 HA antigens would have been as much as 22 years diverged from the pandemic HA (11 years each of independent evolution from the MRCA at 1907 to the pandemic and seasonal lineages evidently circulating in 1918). The 1918 HA would have been 11 years diverged from the putative H1 HA infecting individuals back in 1907. Only a very small proportion of the population would be expected to have been exposed to an HA homologous to the pandemic HA. What was unusual about the 1918 pandemic was not its attack rate (which was typical and consistent with a pre-pandemic interval of H1 HA transmission) but its severity, particularly in young adults.

Possible limitations of the clock model. As mentioned in the main text, it is possible that hidden or reverted host states within the data sets we analyzed may affect our conclusions about the direction of host jumps and the timing of these events. For one thing, as in ref 24, we excluded from our analyses mammalian clades nested within other mammalian clades, which cannot be modeled with the current implementation of the HSLC method. The only such case in the present analysis is the 2009 pH1N1 virus, which originated in swine and moved to humans in or shortly before 2009. We can see no way in which this exclusion would affect our estimates of the local clock rate in the human H1N1 lineage depicted in Figs. 2, S1, and so on: events occurring after 2009 could not have affected inferences based on sequences from earlier years. Moreover, we do not believe there is any chance that the distinct swine and human clades apparent in our trees, from the 1930s onwards, contain hidden host states. There is ample epidemiological evidence that these IAV lineages circulated continuously and separately in each host (with the exception of occasional 'dead-end' jumps that are easy to visualize on the trees and exclude).

More problematic are the longer interior branches in the deeper parts of the tree, prior to the 1930s samples in each lineage. Again, however, there is strong epidemiological evidence that the classical swine influenza virus

circulated in each year from 1918 until the first isolations of the virus in the 1930s, so there seems little doubt that that branch is correctly specified in our analyses. Similarly, most of the stem branch leading up to the seasonal human H1N1 clade is almost certainly correctly specified as human since H1N1 circulated from at least 1918 until the first human isolates of the 1930s.

Least clear is the specification of the branches between 1918 and the common ancestor of the pandemic and seasonal H1 lineages in ~1907. Nevertheless, given the topology of the tree, the most parsimonious scenario is clearly that the human pandemic virus and the seasonal H1N1 lineage shared a common ancestor in humans (Fig. S12A). Other scenarios require one or more additional host jumps (Fig. S12B-F); they also are less consistent with (i) the seroarcheologic evidence entry of H1 antigens into the human populations between about 1896 and 1907 (*1-4, 7-9, 10-12, 15*); (ii) the evidence of several years of mammalian transmission indicated by the high uracil content of the 1918 *HA* gene; (iii) the epidemiological evidence that influenza was new to swine in 1918 (*7, 25*); and (iv) the lack of H3 HA antibodies and lack of protection during the 1968 H3N2 pandemic in those born between ~1900 and 1918, suggesting that a non-H3 IAV subtype circulated in humans in that period (*1-4, 10-13*). These independent lines of evidence suggest (but, we recognize, do not definitively prove) that an H1 HA circulated in humans a decade or so prior to 1918. Moreover, even if one of these more complex scenarios indeed obtained, the timing of the ancestor of the pandemic and seasonal H1 lineages might shift by a few years but the overall pattern of a deep divergence between them would remain, as would the close relationship between the 1918 *HA* and the swine H1N1 *HA* genes (much closer than the pandemic and seasonal *HA* genes are to each other).

It is also worth noting that while the 1930s and later swine H1N1 lineage and the 1930s and later human H1N1 lineage were constrained to be monophyletic in our analyses, both of these assumptions were carefully validated with non-clock trees. Furthermore, the 1918 human sequences were not enforced to be monophyletic with other human strains. Thus, the monophyly constraints imposed are highly unlikely to have biased the results in any substantive way.

Antigenic imprinting and seasonal influenza mortality patterns. The 1968 H3N2 pandemic was notable for its lack of severity in the elderly (*11, 12*). Now that the elderly cohort from 1968—whose childhood exposure to H3 in 1889-1900 provided protection to H3N2 (*12, 13*)—has been supplanted by an elderly cohort primed with H1N1, the pattern has reversed: the same virus is now noteworthy for its severity in the elderly and its relative mildness in younger patients. The current pattern of high H3N2-caused mortality in older adults (*22, 23*) is difficult to reconcile with faltering immunity in the aged: similarly aged adults, evidently primed as children by an H3 HA, had good protection in 1968-69, with no excess mortality. The pattern also seems at odds with the idea that properties intrinsic to this strain of IAV make it particularly dangerous for elderly patients: H3N2 was not inherently virulent in the elderly when it first emerged. While it is possible that intrinsic properties of the virus or weakened immunity in the elderly somehow account for this reversal, cohort immunity shaped by childhood exposure to homosubtypic or heterosubtypic antigens seems more likely in light of our findings. (For H1N1, conversely, the current elderly cohort, who were exposed to H1N1 as children, appear to be protected: we surmise that their childhood exposure to H1N1 reduces the effective number of susceptibles and reduces the frequency of severe postinfluenza complications leading to death.)

Influenza-related hospitalizations are about twice as high in the U.S. during H3N2-dominated seasons (*22*), and H3N2-related all-cause mortalities are more than 10 times higher than for H1N1 (*23*). Deaths among the elderly are disproportionately numerous with H3N2 compared to H1N1. During the 1990s, the estimated annual number of H1N1-related deaths (all cause) among 1-4 year-olds was 34; for H3N2 it was 103 (*23*). For each H1N1 death in the 1-4 age group, there were 57 deaths in those ≥65. However, for H3N2 this ratio was 1:338, almost six times greater. (The numbers of annual H1N1-related and H3N2-related deaths in those ≥65 were 1,954 and 34,866, respectively (*23*), or ~18 H3N2 deaths per H1N1 death in the elderly group.)

This suggests to us a greatly increased risk of death when elderly patients are infected by a seasonal virus with HA and NA glycoproteins heterosubtypic to those of their first exposure in childhood. Conversely, the 5-49 year-old group (who, because of the emergence of an H3N2 virus in 1968 are expected to have relatively good protection against H3N2 compared to those >65, and relatively poor H1N1 protection) accounted for 17.7% of H1N1 related deaths but just 4.2% of H3N2-related deaths. Therefore, as in 1918 and other pandemics, childhood immunity appears to have a strong effect in shaping seasonal influenza mortality patterns. Thompson *et al*. (*23*)

describe H3N2 as the "most virulent of the recently circulating influenza viruses". We hypothesize that this 'virulence' is due largely (or perhaps even completely) to host immunity, not to intrinsic properties of the H3N2 virus.

Antigenic imprinting and H5N1 versus H7N9 mortality patterns. We compared patterns of mortality due to avian-origin H5N1 and H7N9 in different birth year cohorts using H5N1 fatality data from Indonesia in 2005-2005 (*26*) and H7N9 fatality data from China in 2013 (*27*). Although others have noted that H5N1 tends cause severe disease and death primarily in younger age groups and H7N9 in older age groups (*27*), to our knowledge birth year (rather than age at time of infection) it its connection to group 1 versus group 2 HA exposure has not been considered previously.

As depicted in Fig. S15A, 100% of H5N1 fatalities in 2005-2006 in Indonesia occurred among patients born in 1968 or later, after the emergence of H3N2 that occasioned a switch from group 1 HA to group 2 HA. In contrast, 85% of H7N9 fatalities in China in 2013 occurred among those born prior to 1968. We performed a 2x2 Chi-square test for each subtype as follows: we compared the observed number of fatalities in those born prior to 1968 (exposed in childhood to a group 1 HA of either the H1 or H2 subtype) and those born in 1968 and later (exposed in childhood primarily to a group 2 HA of the H3 subtype). We then calculated the expected number of fatalities in each of these two cohorts if cases were evenly distributed, correcting for the different population size of each cohort by using census data for Indonesia in 2006 or China in 2013 (*28*).

For H5N1, the observed and expected post-1967 fatalities were 41 and 28, respectively, and the observed and expected pre-1968 fatalities were 0 and 13, respectively (Chi-square=15.4, df=1, two-tailed $P<0.0001$). For H7N9, the observed and expected post-1967 fatalities were 5 and 23, respectively, and the observed and expected pre-1968 fatalities were 29 and 11, respectively (Chi-square=19.7, df=1, two-tailed $P<0.0001$). Note that the age pattern is reversed for the different viruses, consistent with initial childhood exposure to a 'group-mismatched' HA greatly increasing mortality risk, and heterosubtypic but 'group-matched' HA providing near complete protection from fatal disease upon exposure to a novel, avian-origin influenza A virus. Combining both H5N1 and H7N9 fatalities suggests that ~93% of fatalities have occurred in individuals exposed to an avian virus with an HA mismatched to their childhood HA group.

This can be seen in Fig. S15B, in which census data (*28*) are used to account for the proportion of the total population (in each country) contained in each 10-year birth year bin. (Values above 0 indicate a greater number of fatalities than if the fatalities were distributed in proportion to the underlying demographic distribution.) The distributions for the two viruses are near-mirror images, centered on the switch from group 1 HA to group 2 HA in 1968. Moreover, much of the ~15% mortality in the right tail of the H7N9 distribution (in birth years of 1968 or later) is plausibly due to the re-emergence of H1 in 1977, which we hypothesize may make a minority of those born after 1977 (in particular those born close to the so-called 'pandemic' in 1977) more susceptible to severe disease when infected with a group-mismatched (H7) HA virus. If this is correct, then the co-circulation of H3N2 and H1N1 since 1977 may shape severity outcomes in those born from ~1977 onward such that those initially exposed the group 1 virus (H1N1) are more susceptible later to H3N2 and H7N9 and those initially exposed to the group 2 virus (H3N2) are more susceptible later to H1N1 and H5N1. It may be possible to test this hypothesis with epidemiologic or immunologic approaches.

Supplementary References

1. Masurel N, Heijtink RA (1983) Recycling of H1N1 influenza A virus in man—a haemagglutinin antibody study. *J Hyg (Lond)* 90(3):397-402.
2. Masurel N (1969) Relation between Hong Kong virus and former human A2 isolates and the A-EQU12 virus in human sera collected before 1957. *Lancet* 1(7601):907-910.
3. Masurel N, Mulder J (1966) Studies on the content of antibodies for equine influenza viruses in human sera. *Bull World Health Organ* 34(6):885-893.
4. Davenport FM, Minuse E, Hennessy AV, Francis T (1969) Interpretations of influenza antibody patterns of man. *Bull World Health Organ* 41(3):453-460.

5. Dowdle WR (1999) Influenza A virus recycling revisited. *Bull World Health Organ* 77(10):820-828.
6. Francis T (1960) On the doctrine of original antigenic sin. *Proc Am Philos Soc* 104(6):572-578.
7. Shope RE (1936) The incidence of neutralizing antibodies for swine influenza virus in the sera of human beings of different ages. *J Exp Med* 63(5):669-684.
8. Francis T, Magill TP (1936) The incidence of neutralizing antibodies for human influenza virus in the serum of human individuals of different ages. *J Exp Med* 63(5):655-668.
9. Rekart M, Rupnik K, Cesario TC, Tilles JG (1982) Prevalence of hemagglutination inhibition antibody to current strains of the H3N2 and H1N1 subtypes of influenza A virus in sera collected from the elderly in 1976. *Am J Epidemiol* 115(4):587-597.
10. Kendal AP, Minuse E, Maassab HF, Hennessy AV, Davenport FM (1973) Influenza neuraminidase antibody patterns of man. *Am J Epidemiol* 98(2):96-103.
11. Simonsen L, Reichert TA, Miller C (2004) The virtues of antigenic sin: consequences of pandemic recycling on influenza-associated mortality. *International Congress Series* 1263:791-794.
12. Schoenbaum SC, Coleman MT, Dowdle WR, Mostow SR (1976) Epidemiology of influenza in the elderly: evidence of virus recycling. *Am J Epidemiol* 103(2):166-173.
13. Davenport FM, Minuse E, Hennessy AV, Francis T (1969) Interpretations of influenza antibody patterns of man. *Bull World Health Organ* 41(3):453-460.
14. Masurel N (1976) Swine influenza virus and the recycling of influenza-A viruses in man. *Lancet* 2(7979):244-247.
15. Bodewes R, et al. (2011) Prevalence of antibodies against seasonal influenza A and B viruses in children in Netherlands. *Clin Vaccine Immunol* 18(3):469-476.
16. Andrewes CH, Laidlaw PP, Smith W (1935) Influenza: Observations on the recovery of virus from man and on the antibody content of human sera. *Br J Exp Pathol* 16(6):566–582.
17. Davenport FM, Hennessy AV, Drescher J, Mulder J, Francis T (1964) Further observations on the relevance of serologic recapitulations of human infection with influenza viruses. *J Exp Med* 120(6):1087-1097.
18. Law J (1874) *Report of the Commissioner of Agriculture, 1872: Influenza in Horses* (Government Printing Office, Washington).
19. Patterson KD (1986) *Pandemic influenza, 1700-1900: A study in historical epidemiology*. (Rowan & Littlefield, Totowa, N. J.).
20. Fan J, et al. (2004) Preclinical study of influenza virus A M2 peptide conjugate vaccines in mice, ferrets, and rhesus monkeys. *Vaccine* 22 (23-24):2993-3003.
21. Frost WH (1919) The epidemiology of influenza, 1919. *Public Health Rep* 34 (33):1823–1861.
22. Thompson WW, et al. (2004) Influenza-associated hospitalizations in the United States. *JAMA* 292(11):1333-1340.
23. Thompson WW, et al. (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. *JAMA* 289(2):179-186.
24. Worobey M, Han GZ, Rambaut A (2014) A synchronized sweep of the internal genes of modern avian influenza virus. *Nature* doi:10.1038/nature13016.
25. Koen JS (1919) A practical method for the field diagnoses of swine diseases. *Am J Vet Med* 14(9):468-470.
26. Sedyaningsih ER, et al. (2007) Epidemiology of cases of H5N1 virus infection in Indonesia, July 2005-June 2006. *J Inf Dis* 196(4):522-527.
27. Dudley JP, MacKay IM (2013) Age-specific and sex-specific morbidity and mortality from avian influenza A(H7N9). *J Clin Virol* 58(3):568– 570.
28. http://www.census.gov/population/international/data/idb/informationGateway.php

Supplementary Figures S1 to S15

Fig. S1. H1 MCC tree and circulating diversity in 1918. (**A**) H1 MCC tree (the full version of the tree depicted in Fig. 2). Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk. (**B**) Years of genetic diversity circulating in *HA* in 1918 (i.e. the time to the most recent common ancestor of the pandemic and seasonal *HA* lineages sampled at each time point) compared to circulating diversity in other years (both for seasonal H1N1 and for later pandemics). Circles indicate median node dates and gray rectangles indicate 95% CIs.

A

B

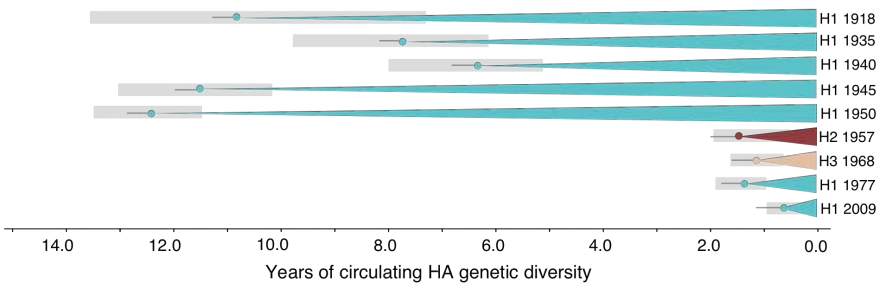

H1 1918
H1 1935
H1 1940
H1 1945
H1 1950
H2 1957
H3 1968
H1 1977
H1 2009
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
Years of circulating HA genetic diversity

Fig. S2. H1 MCC tree (1930s laboratory strains and 1918 short sequence fragments removed). Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk.

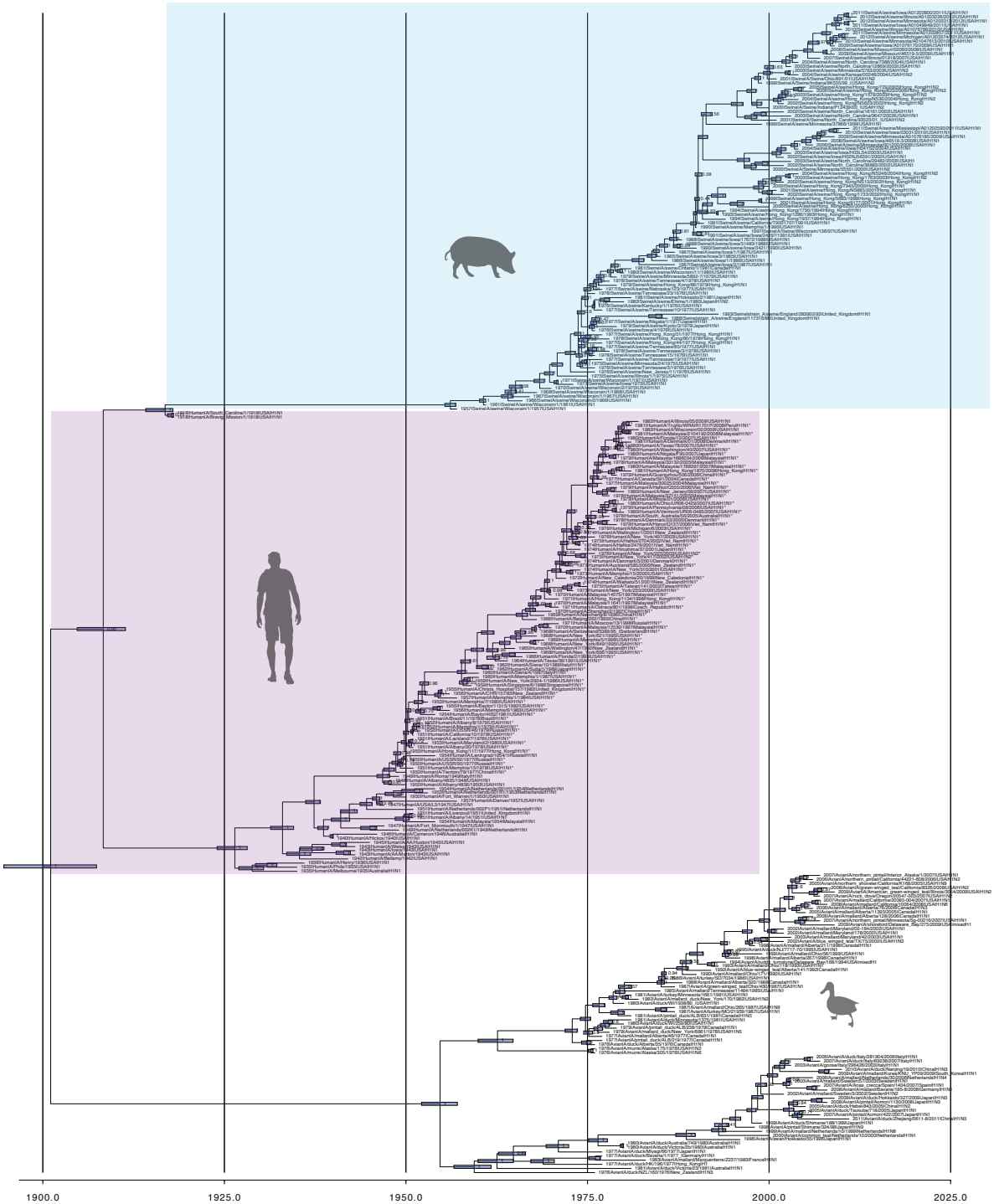


Fig. S3. H1 MCC tree (1930s laboratory strains and 1918 short sequence fragments removed, maximum of one sequence/year/host, post-1957 human H1N1 sequences removed). Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown.

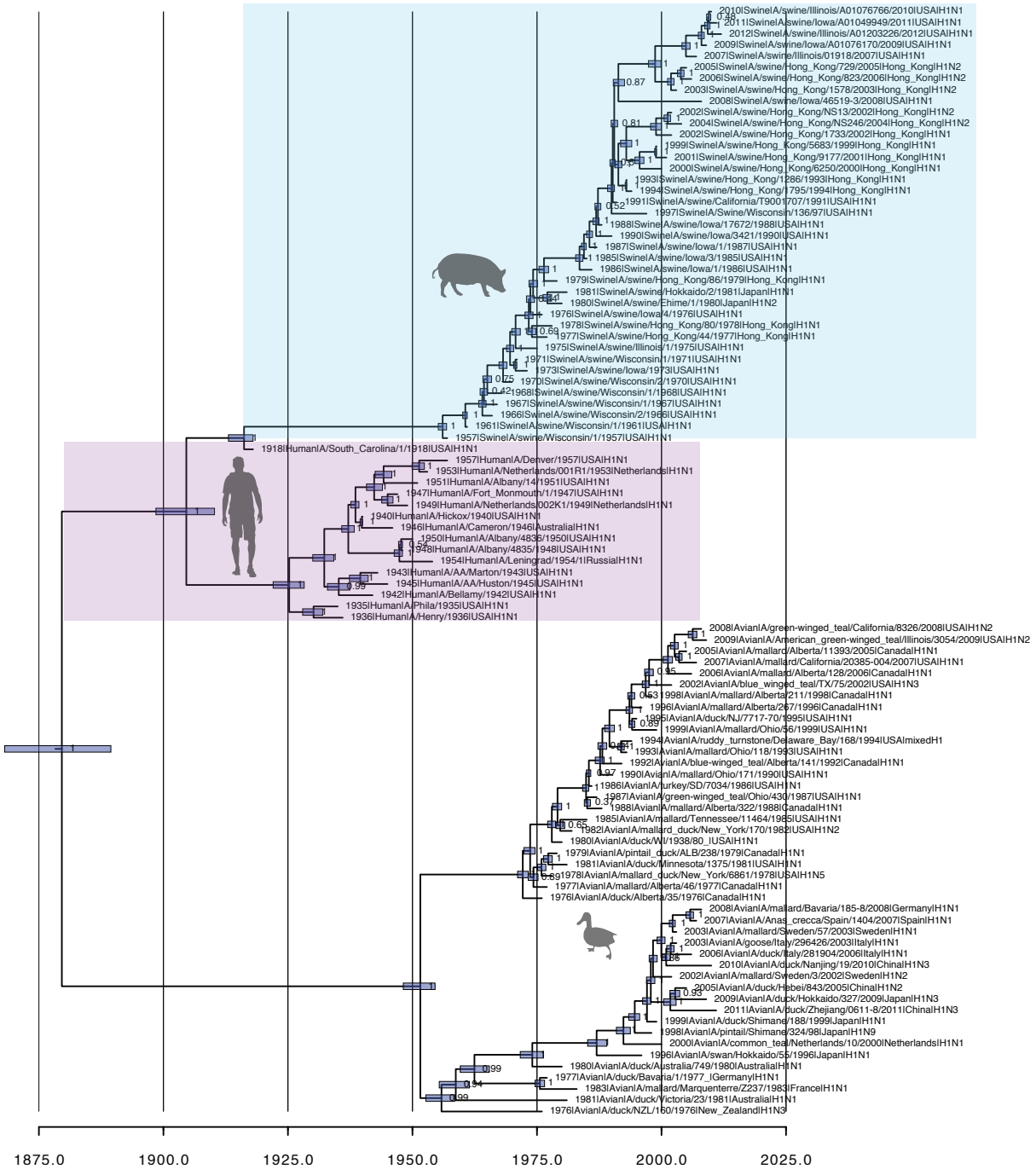


Fig. S4. H1 MCC tree (HA2/stalk domain only). Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk.

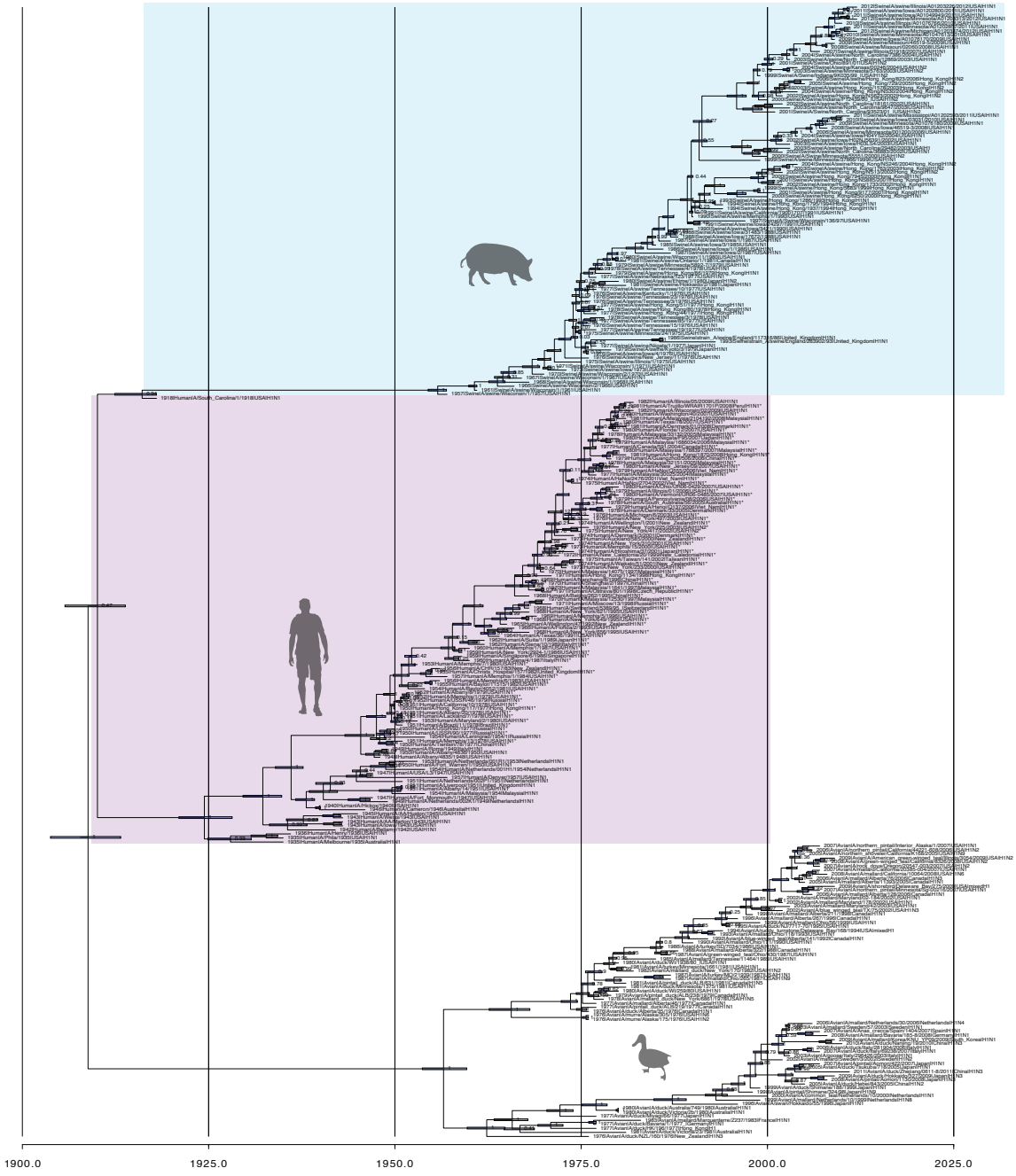


Fig. S5. Branch-site REL analyses to test for episodic diversifying selection. The branches are colored to depict the proportion of substitutions along each branch that are under purifying selection (with dN/dS < 1: blue), the proportion evolving neutrally (with dN/dS = 1: gray), and the proportion under diversifying selection (with dN/dS > 1: red). Almost every site in every branch evidently evolved under purifying selection, and none of the branches exhibited significant evidence of episodic diversifying selection.

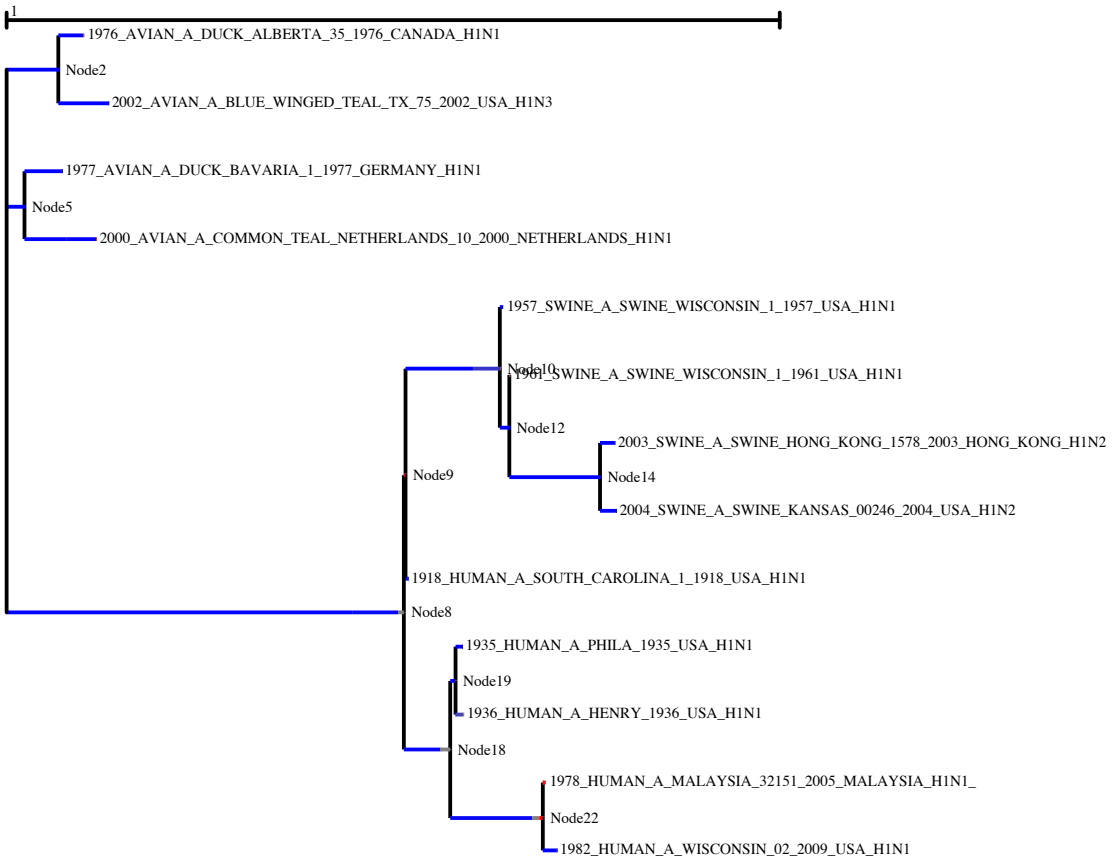
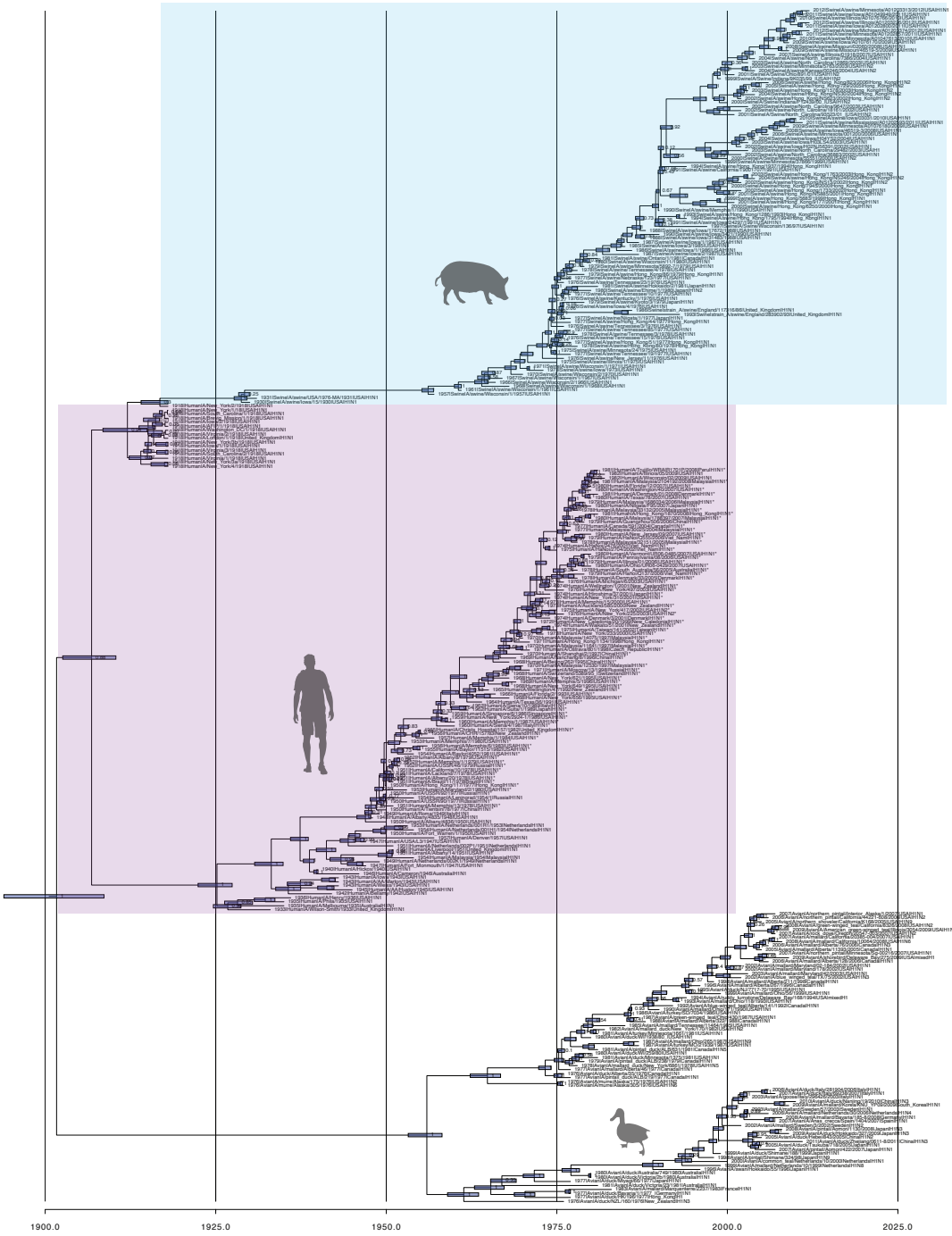
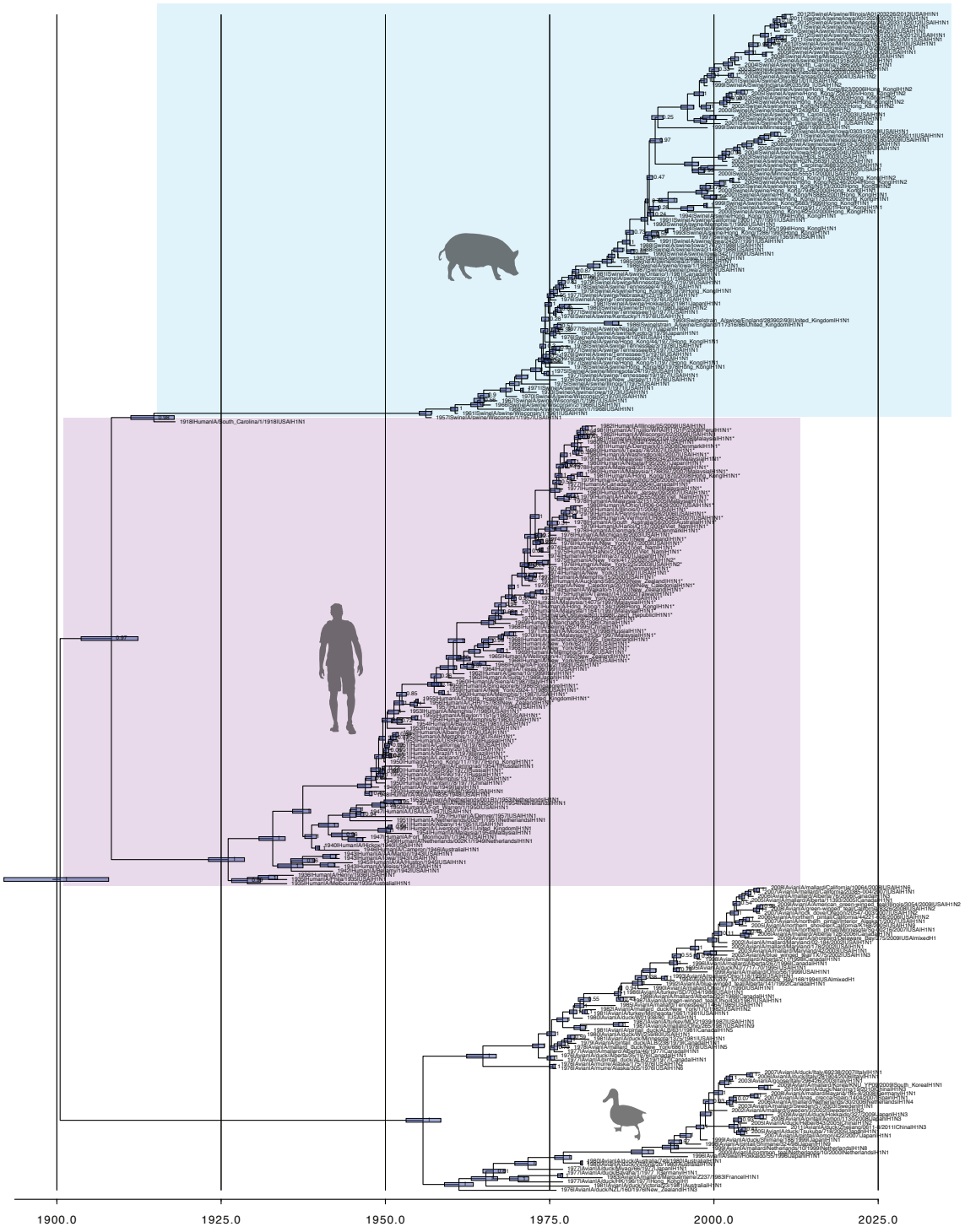


Fig. S6. H1 MCC tree (**A**) 1st and 2nd codon position sites removed (i.e., 3rd sites only). (**B**) 3rd sites only, plus 1930s laboratory strains and 1918 short sequence fragments removed. (**C**) Same as **B** but including only 3rd position sites at which substitutions were synonymous (silent sites). In all cases branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk.

A



B



C

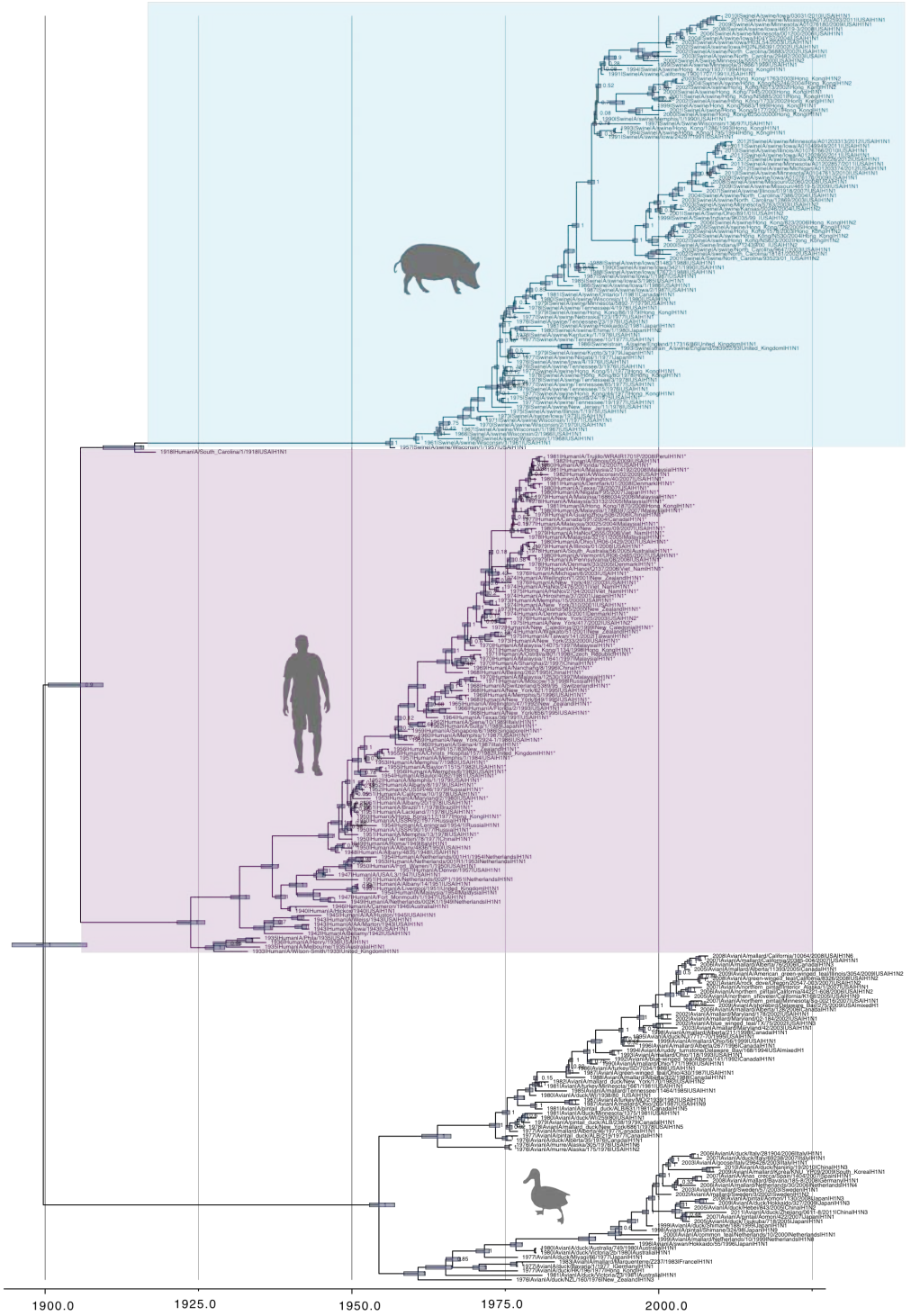


Fig. S7. Uracil (U) content patterns. The 95% CI of avian U content is shown for each segment with a gray rectangle. U content versus year of sampling is shown by magenta diamonds for the 1918 human H1N1 sequences of (**A**) N1 *NA*, and (**B**) H1 *HA*. The *P* value in panel **B** reflects a test of whether the U content of the 1918 *HA* is significantly higher than the avian range based on the rate of change observed in the H3N2 lineage.

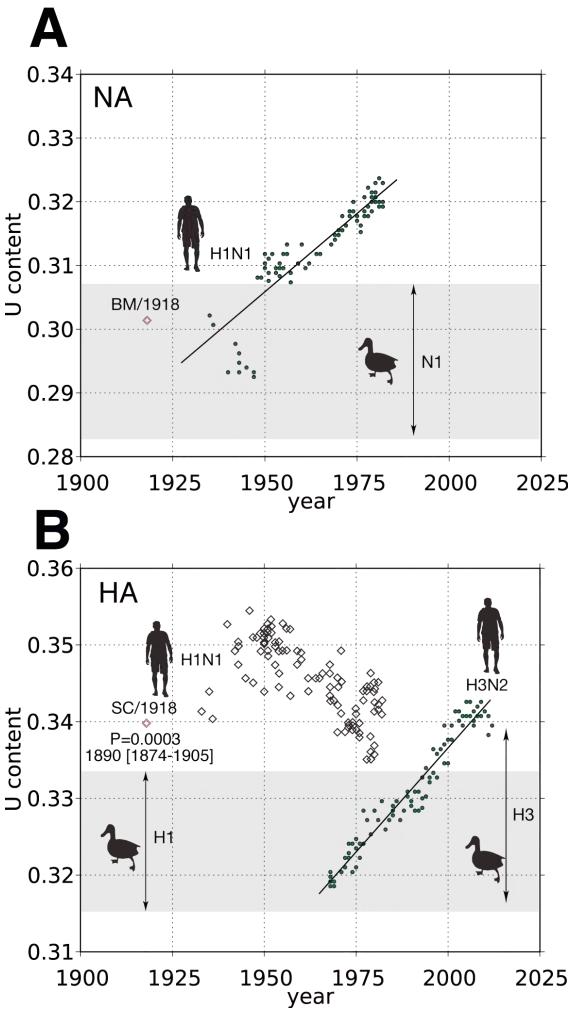


Fig. S8. N1 MCC tree. Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1N1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk. The 95% CI for the within-human H1 diversity is shown for comparison.

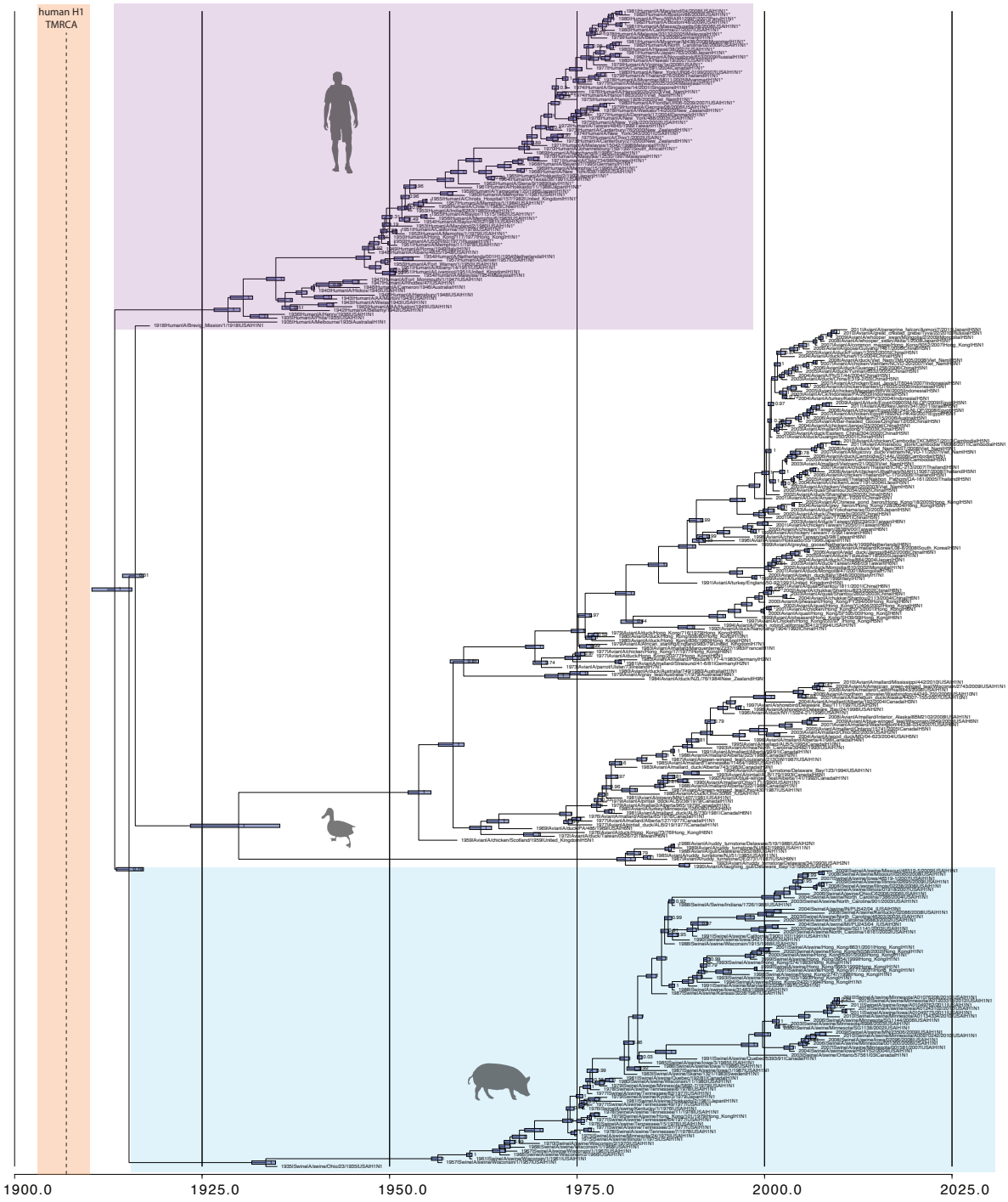


Fig. S9. *PB1* MCC tree. Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1N1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk. The 95% CI for the within-human H1 diversity is shown for comparison.

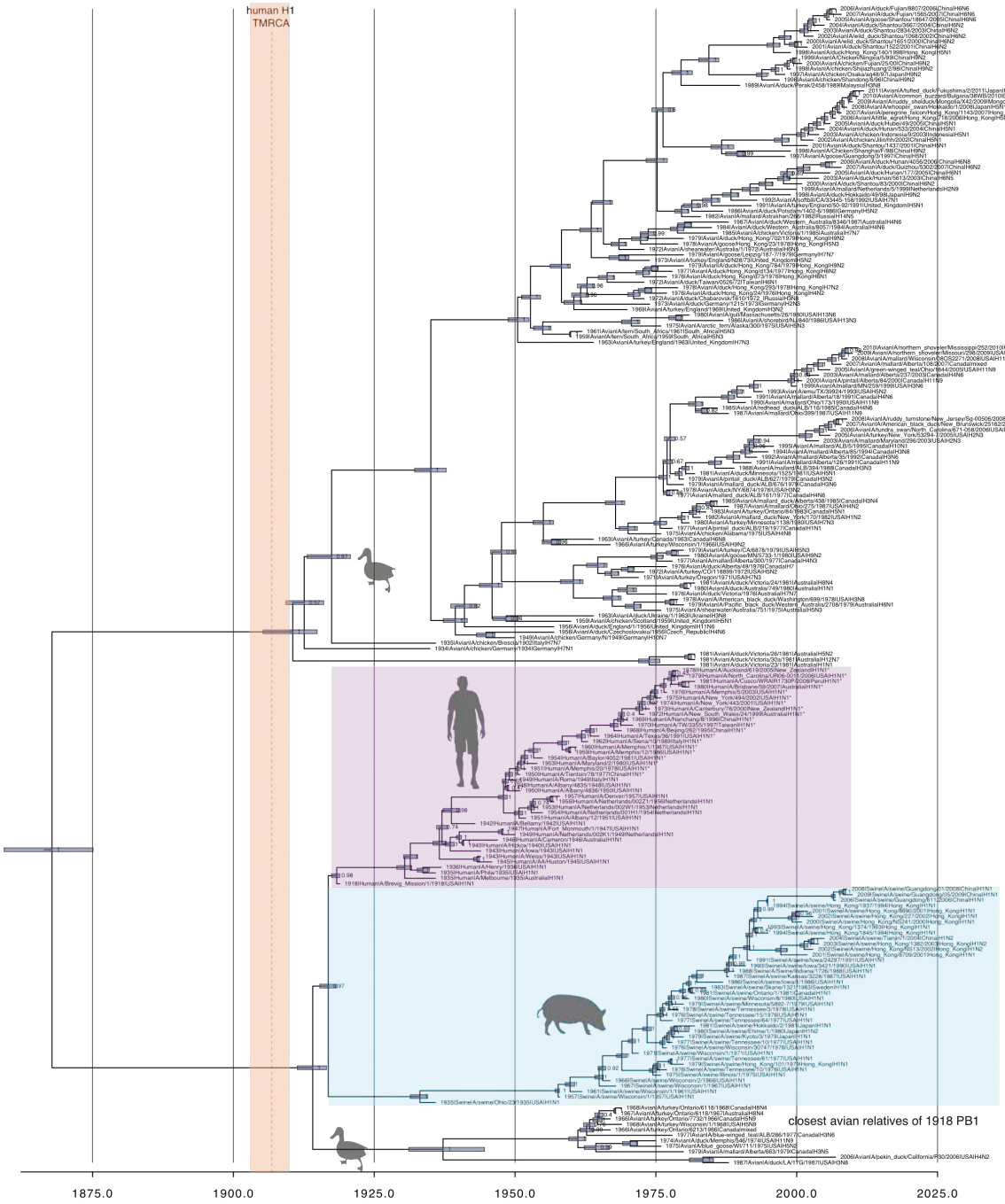


Fig. S10. H1 in humans predates all other genes in the 1918 pandemic virus. The gray bar indicates the window of overlap among the 95% CIs of the TMRCA of human and swine H1N1 for the different segments. The TMRCA of human H1 significantly predates that of each of the other genes. The TMRCA of the human pandemic and season lineages also predates the TMRCA of the classical swine/pandemic human H1 node.

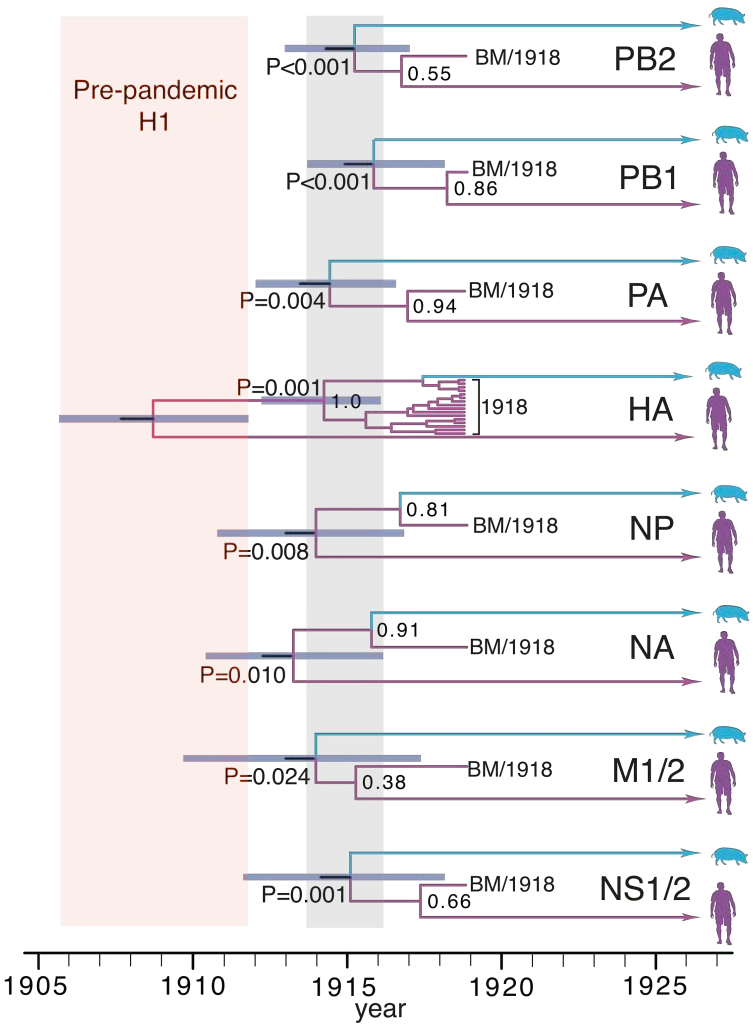
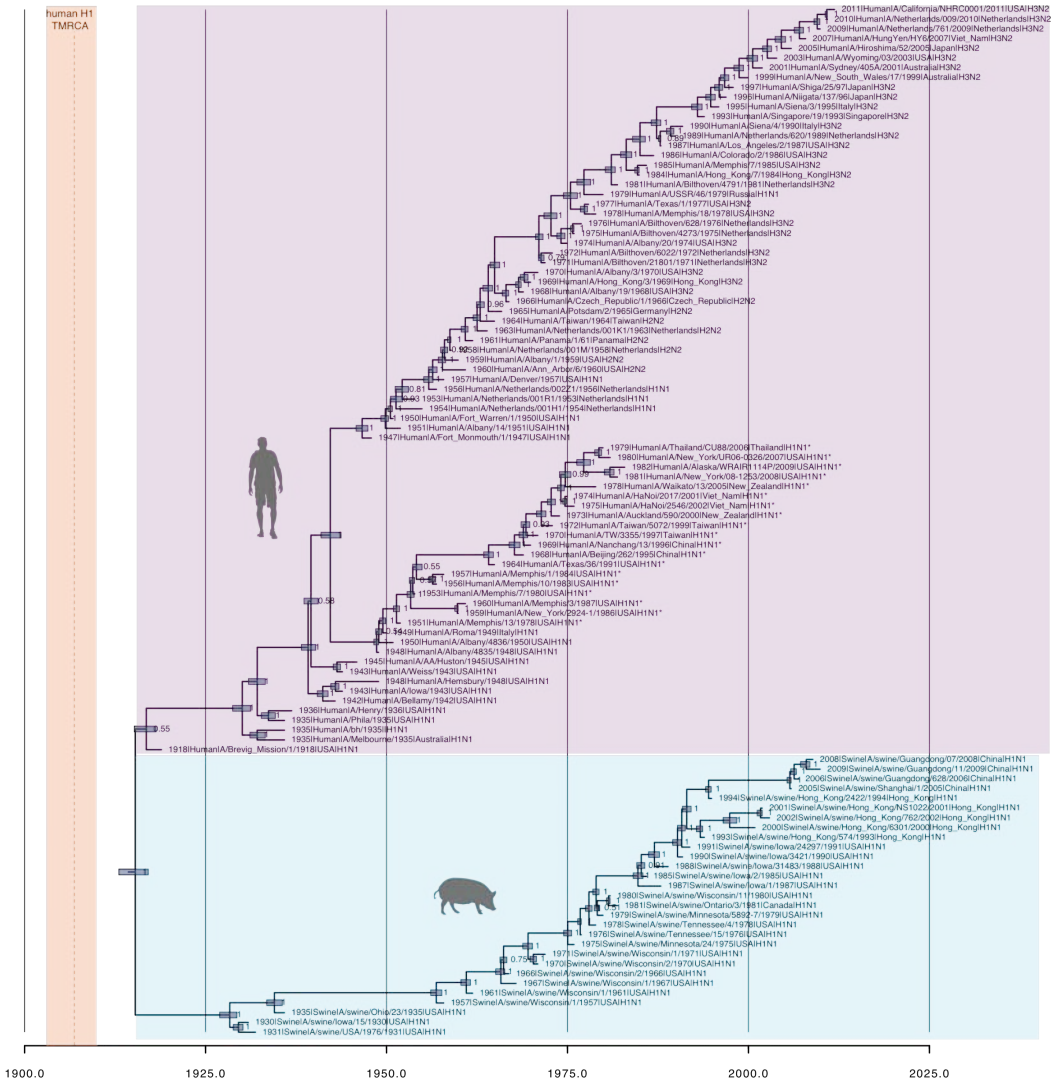
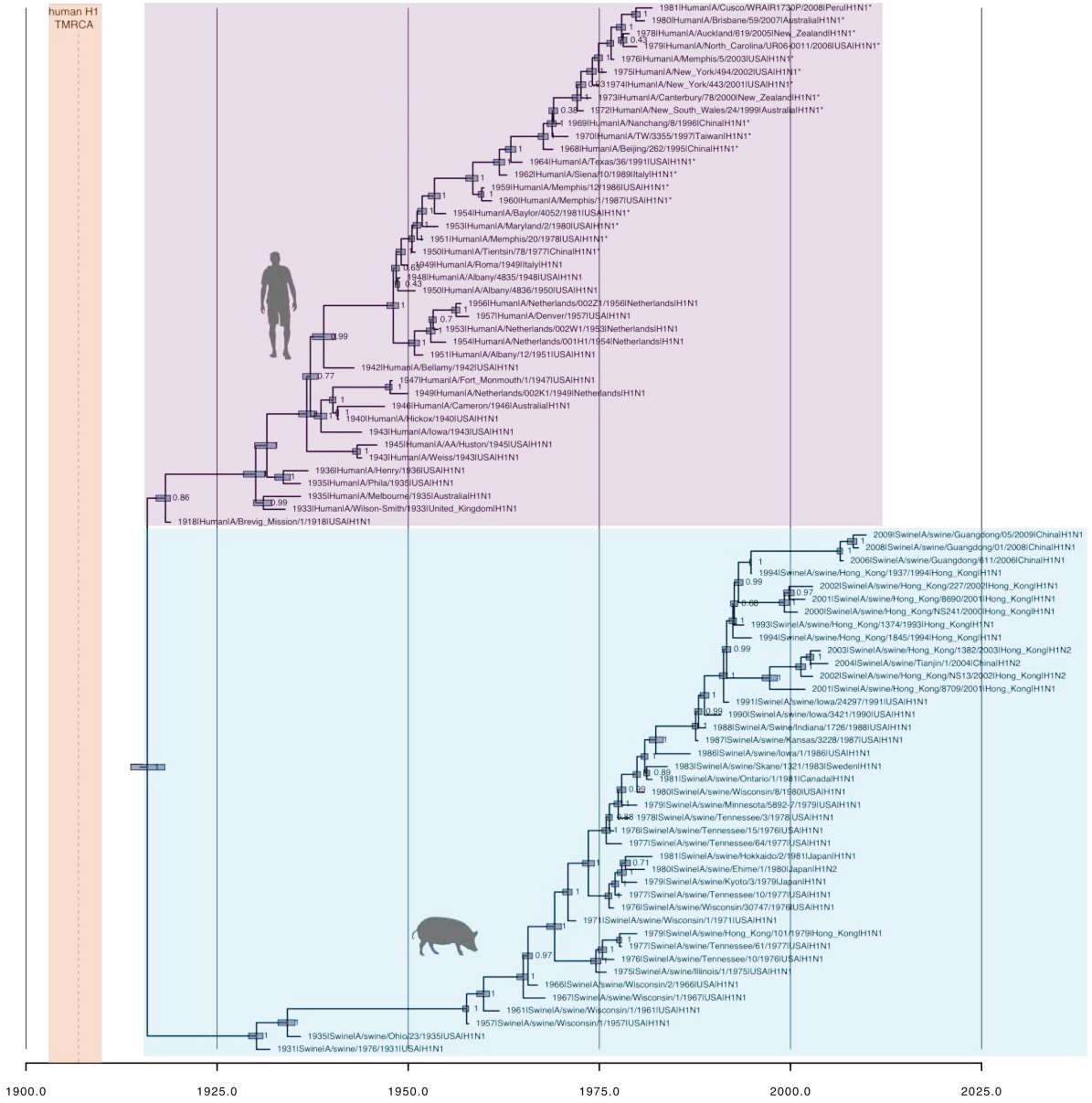


Fig. S11. Human and swine IAV MCC trees for all 8 segments. (**A**) through (**H**) respectively: *PB2*, *PB1*, *PA*, *HA*, *NP*, *NA*, *M1/2*, *NS1/2*. Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk. The 95% CI for the within-human H1 diversity is shown for comparison.

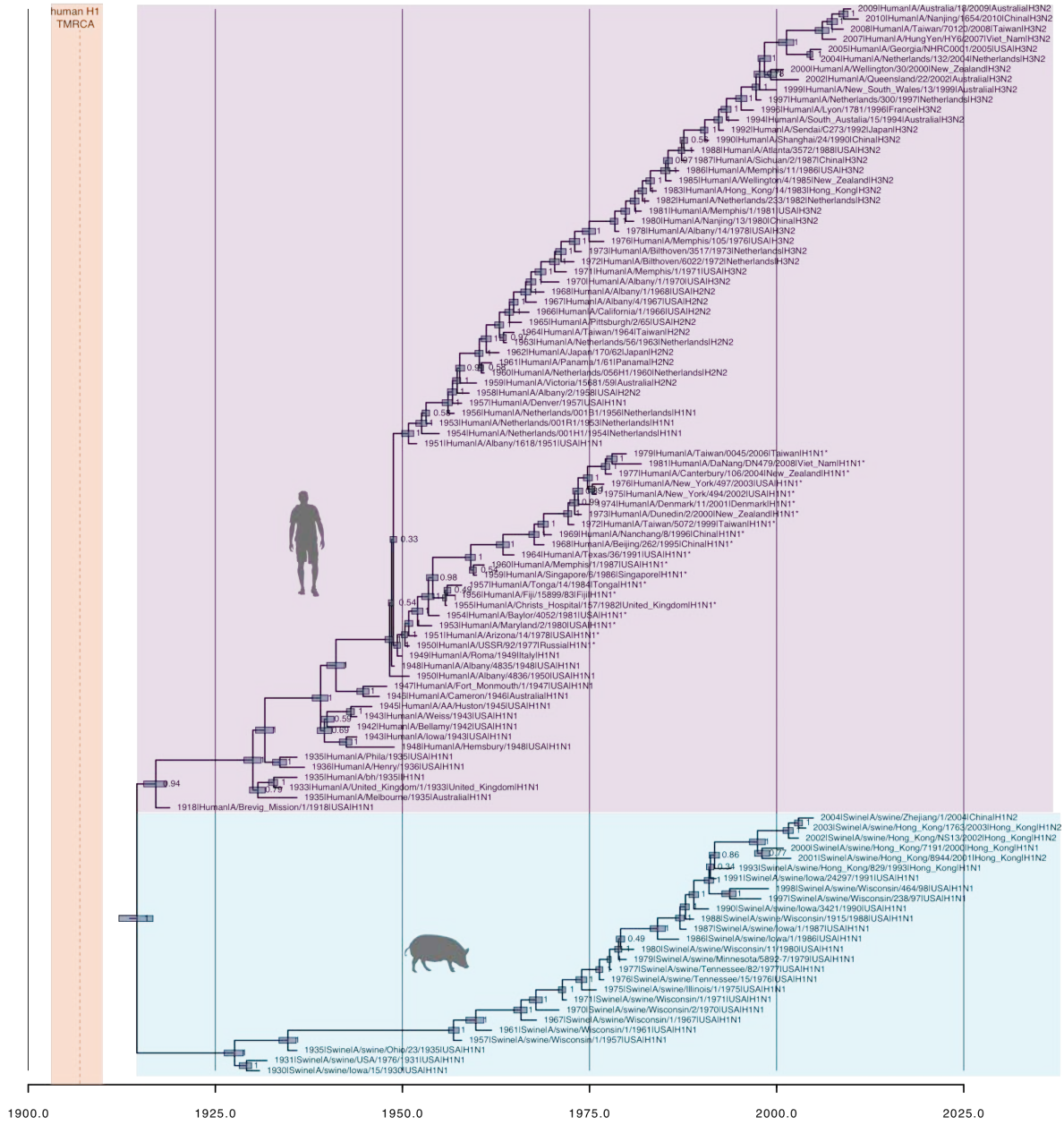
A *PB2*



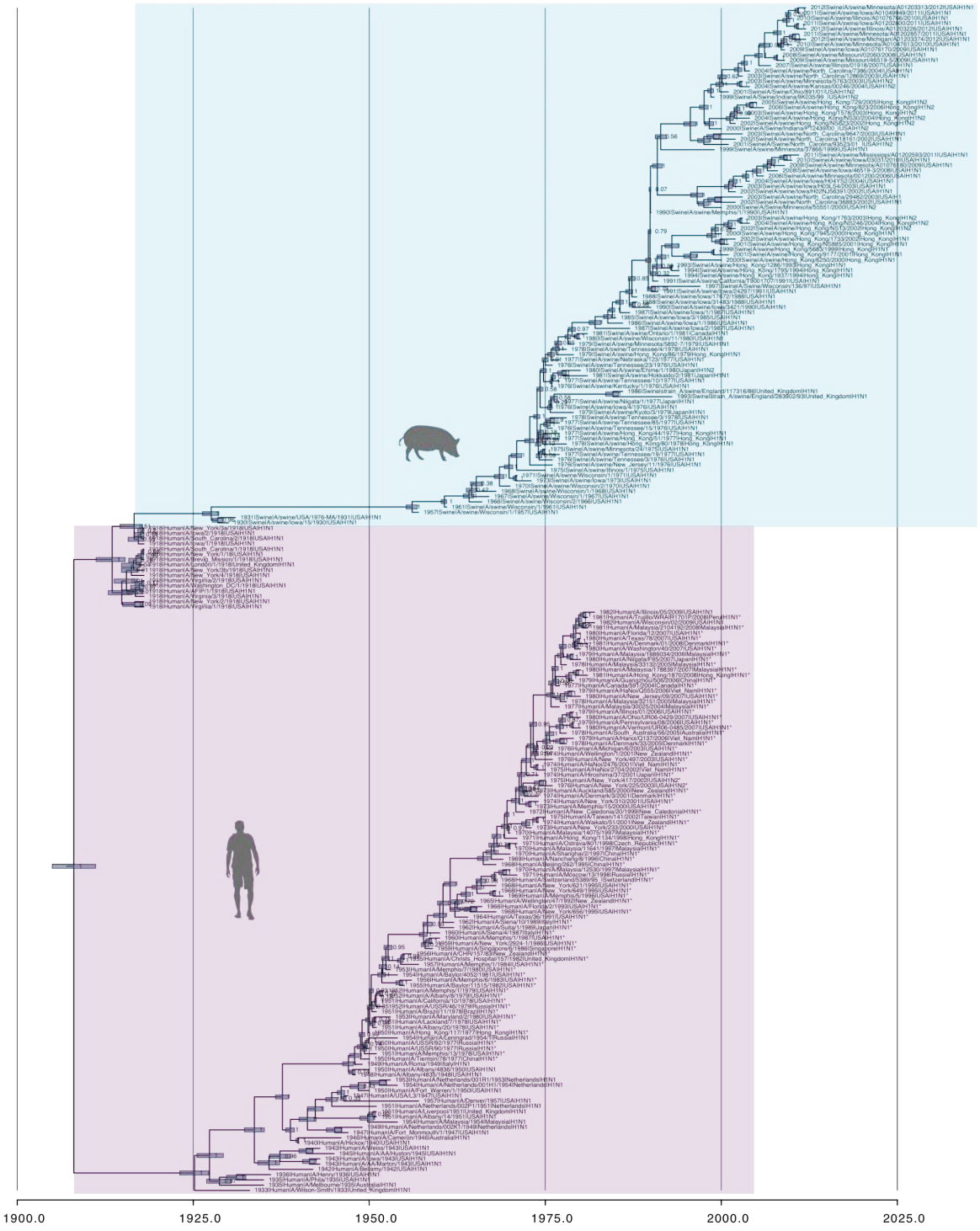
B *PB1*


human H1 TMRCA
1981|HumanIA/Cusco/WRAIR1730P/2008|PeruIH1N1*
1980|HumanIA/Brisbane/59/2007|AustralialH1N1*
1978|HumanIA/Auckland/619/2005|New_ZealandIH1N1*
1979|HumanIA/North_Carolina/UR06-0011/2006|USAIH1N1*
1976|HumanIA/Memphis/5/2003|USAIH1N1*
1975|HumanIA/New_York/494/2002|USAIH1N1*
1974|HumanIA/New_York/443/2001|USAIH1N1*
1973|HumanIA/Canterbury/78/2000|New_ZealandIH1N1*
1972|HumanIA/New_South_Wales/24/1999|AustralialH1N1*
1969|HumanIA/Nanchang/8/1996|ChinaIH1N1*
1970|HumanIA/TW/3355/1997|TaiwanIH1N1*
1968|HumanIA/Beijing/262/1995|ChinaIH1N1*
1964|HumanIA/Texas/36/1991|USAIH1N1*
1962|HumanIA/Siena/10/1989|ItalyIH1N1*
1959|HumanIA/Memphis/12/1986|USAIH1N1*
1960|HumanIA/Memphis/1/1987|USAIH1N1*
1954|HumanIA/Baylor/4052/1981|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1951|HumanIA/Memphis/20/1978|USAIH1N1*
1950|HumanIA/Tientsin/78/1977|ChinaIH1N1*
1949|HumanIA/Roma/1949|ItalyIH1N1
1948|HumanIA/Albany/4835/1948|USAIH1N1
1950|HumanIA/Albany/4836/1950|USAIH1N1
1956|HumanIA/Netherlands/002Z1/1956|NetherlandsIH1N1
1957|HumanIA/Denver/1957|USAIH1N1
1953|HumanIA/Netherlands/002W1/1953|NetherlandsIH1N1
1954|HumanIA/Netherlands/001H1/1954|NetherlandsIH1N1
1951|HumanIA/Albany/12/1951|USAIH1N1
1942|HumanIA/Bellamy/1942|USAIH1N1
1947|HumanIA/Fort_Monmouth/1/1947|USAIH1N1
1949|HumanIA/Netherlands/002K1/1949|NetherlandsIH1N1
1946|HumanIA/Cameron/1946|AustralialH1N1
1940|HumanIA/Hickox/1940|USAIH1N1
1943|HumanIA/Iowa/1943|USAIH1N1
1945|HumanIA/AA/Huston/1945|USAIH1N1
1943|HumanIA/Weiss/1943|USAIH1N1
1936|HumanIA/Henry/1936|USAIH1N1
1935|HumanIA/Phila/1935|USAIH1N1
1935|HumanIA/Melbourne/1935|AustralialH1N1
1933|HumanIA/Wilson-Smith/1933|United_KingdomIH1N1
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1
2009|SwineIA/swine/Guangdong/05/2009|ChinaIH1N1
2008|SwineIA/swine/Guangdong/01/2008|ChinaIH1N1
2006|SwineIA/swine/Guangdong/611/2006|ChinaIH1N1
1994|SwineIA/swine/Hong_Kong/1937/1994|Hong_KongIH1N1
2002|SwineIA/swine/Hong_Kong/227/2002|Hong_KongIH1N1
2001|SwineIA/swine/Hong_Kong/8690/2001|Hong_KongIH1N1
2000|SwineIA/swine/Hong_Kong/NS241/2000|Hong_KongIH1N1
1993|SwineIA/swine/Hong_Kong/1374/1993|Hong_KongIH1N1
1994|SwineIA/swine/Hong_Kong/1845/1994|Hong_KongIH1N1
2003|SwineIA/swine/Hong_Kong/1382/2003|Hong_KongIH1N2
2004|SwineIA/swine/Tianjin/1/2004|ChinaIH1N2
2002|SwineIA/swine/Hong_Kong/NS13/2002|Hong_KongIH1N2
2001|SwineIA/swine/Hong_Kong/8709/2001|Hong_KongIH1N1
1991|SwineIA/swine/Iowa/24297/1991|USAIH1N1
1990|SwineIA/swine/Iowa/3421/1990|USAIH1N1
1988|SwineIA/Swine/Indiana/1726/1988|USAIH1N1
1987|SwineIA/swine/Kansas/3228/1987|USAIH1N1
1986|SwineIA/swine/Iowa/1/1986|USAIH1N1
1983|SwineIA/swine/Skane/1321/1983|SwedenIH1N1
1981|SwineIA/swine/Ontario/1/1981|CanadaIH1N1
1980|SwineIA/swine/Wisconsin/8/1980|USAIH1N1
1979|SwineIA/swine/Minnesota/5892-7/1979|USAIH1N1
1978|SwineIA/swine/Tennessee/3/1978|USAIH1N1
1976|SwineIA/swine/Tennessee/15/1976|USAIH1N1
1977|SwineIA/swine/Tennessee/64/1977|USAIH1N1
1981|SwineIA/swine/Hokkaido/2/1981|JapanIH1N1
1980|SwineIA/swine/Ehime/1/1980|JapanIH1N2
1979|SwineIA/swine/Kyoto/3/1979|JapanIH1N1
1977|SwineIA/swine/Tennessee/10/1977|USAIH1N1
1976|SwineIA/swine/Wisconsin/30747/1976|USAIH1N1
1971|SwineIA/swine/Wisconsin/1/1971|USAIH1N1
1979|SwineIA/swine/Hong_Kong/101/1979|Hong_KongIH1N1
1977|SwineIA/swine/Tennessee/61/1977|USAIH1N1
1976|SwineIA/swine/Tennessee/10/1976|USAIH1N1
1975|SwineIA/swine/Illinois/1/1975|USAIH1N1
1966|SwineIA/swine/Wisconsin/2/1966|USAIH1N1
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1
1931|SwineIA/swine/1976/1931|USAIH1N1
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

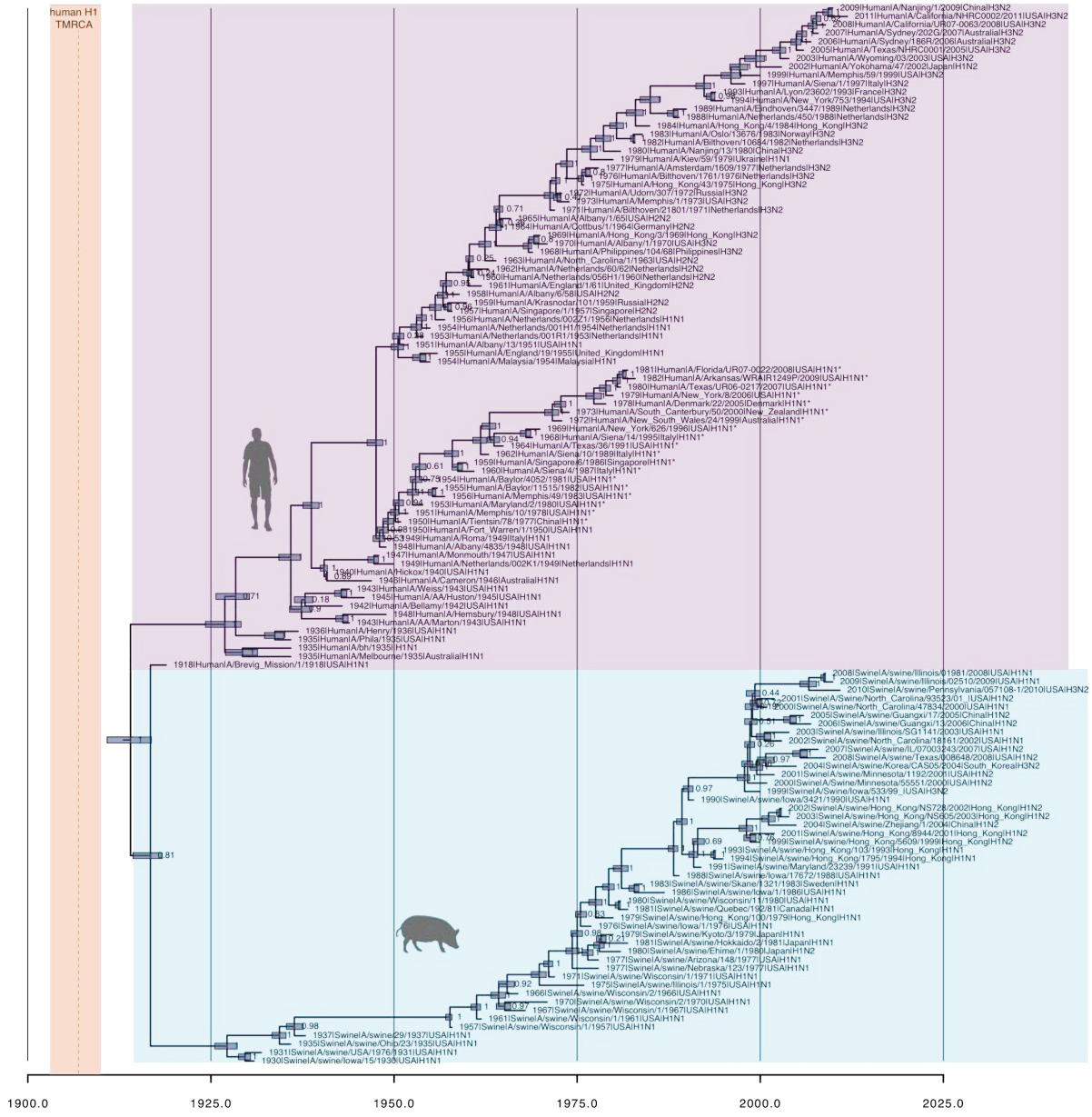
C *PA*


human H1 TMRCA
2009|HumanIA/Australia/18/2009|AustralialH3N2
2010|HumanIA/Nanjing/1654/2010|ChinaIH3N2
2008|HumanIA/Taiwan/7012D/2008|TaiwanIH3N2
2007|HumanIA/HungYen/HY62/2007|Viet_NamIH3N2
2005|HumanIA/Georgia/NHRC001/2005|USAIH3N2
2004|HumanIA/Netherlands/132/2004|NetherlandsH3N2
2000|HumanIA/Wellington/30/2000|New_ZealandIH3N2
2002|HumanIA/Queensland/22/2002|AustralialH3N2
1999|HumanIA/New_South_Wales/13/1999|AustralialH3N2
1997|HumanIA/Netherlands/300/1997|NetherlandsH3N2
1996|HumanIA/Lyon/1781/1996|FranceIH3N2
1994|HumanIA/South_Australia/15/1994|AustralialH3N2
1992|HumanIA/Sendai/C273/1992|JapanIH3N2
1990|HumanIA/Shanghai/24/1990|ChinaIH3N2
1988|HumanIA/Atlanta/3572/1988|USAIH3N2
1987|HumanIA/Sichuan/2/1987|ChinaIH3N2
1986|HumanIA/Memphis/11/1986|USAIH3N2
1985|HumanIA/Wellington/4/1985|New_ZealandIH3N2
1983|HumanIA/Hong_Kong/14/1983|Hong_KongIH3N2
1982|HumanIA/Netherlands/233/1982|NetherlandsIH3N2
1981|HumanIA/Memphis/1/1981|USAIH3N2
1980|HumanIA/Nanjing/13/1980|ChinaIH3N2
1978|HumanIA/Albany/14/1978|USAIH3N2
1976|HumanIA/Memphis/105/1976|USAIH3N2
1973|HumanIA/Bilthoven/3517/1973|NetherlandsIH3N2
1972|HumanIA/Bilthoven/6022/1972|NetherlandsIH3N2
1971|HumanIA/Memphis/1/1971|USAIH3N2
1970|HumanIA/Albany/1/1970|USAIH3N2
1968|HumanIA/Albany/1/1968|USAIH2N2
1967|HumanIA/Albany/4/1967|USAIH2N2
1966|HumanIA/California/1/1966|USAIH2N2
1965|HumanIA/Pittsburgh/2/65|USAIH2N2
1964|HumanIA/Taiwan/1964|TaiwanIH2N2
1963|HumanIA/Netherlands/56/1963|NetherlandsIH2N2
1962|HumanIA/Japan/170/62|JapanIH2N2
1961|HumanIA/Panama/1/61|PanamaIH2N2
1960|HumanIA/Netherlands/056H1/1960|NetherlandsIH2N2
1959|HumanIA/Victoria/15681/59|AustralialH2N2
1958|HumanIA/Albany/2/1958|USAIH2N2
1957|HumanIA/Denver/1957|USAIH1N1
1956|HumanIA/Netherlands/001B1/1956|NetherlandsIH1N1
1953|HumanIA/Netherlands/001R1/1953|NetherlandsIH1N1
1954|HumanIA/Netherlands/001H1/1954|NetherlandsIH1N1
1951|HumanIA/Albany/1618/1951|USAIH1N1
1979|HumanIA/Taiwan/0045/2006|TaiwanIH1N1*
1981|HumanIA/DaNang/DN479/2008|Viet_NamIH1N1*
1977|HumanIA/Canterbury/106/2004|New_ZealandIH1N1*
1976|HumanIA/New_York/497/2003|USAIH1N1*
1975|HumanIA/New_York/494/2002|USAIH1N1*
1974|HumanIA/Denmark/11/2001|DenmarkIH1N1*
1973|HumanIA/Dunedin/2/2000|New_ZealandIH1N1*
1972|HumanIA/Taiwan/5072/1999|TaiwanIH1N1*
1969|HumanIA/Nanchang/8/1996|ChinaIH1N1*
1968|HumanIA/Beijing/262/1995|ChinaIH1N1*
1964|HumanIA/Texas/36/1991|USAIH1N1*
1960|HumanIA/Memphis/1/1987|USAIH1N1*
1959|HumanIA/Singapore/6/1986|SingaporeIH1N1*
1957|HumanIA/Tonga/14/1984|TongaIH1N1*
1956|HumanIA/Fiji/15899/83|FijiIH1N1*
1955|HumanIA/Christs_Hospital/157/1982|United_KingdomIH1N1*
1954|HumanIA/Baylor/4052/1981|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1951|HumanIA/Arizona/14/1978|USAIH1N1*
1950|HumanIA/USSR/92/1977|RussiaIH1N1*
1949|HumanIA/Roma/1949|ItalyIH1N1
1948|HumanIA/Albany/4835/1948|USAIH1N1
1950|HumanIA/Albany/4836/1950|USAIH1N1
1947|HumanIA/Fort_Monmouth/1/1947|USAIH1N1
1945|HumanIA/Cameron/1946|AustralialH1N1
1945|HumanIA/AA/Huston/1945|USAIH1N1
1943|HumanIA/Weiss/1943|USAIH1N1
1942|HumanIA/Bellamy/1942|USAIH1N1
1943|HumanIA/Iowa/1943|USAIH1N1
1948|HumanIA/Hemsbury/1948|USAIH1N1
1935|HumanIA/Phila/1935|USAIH1N1
1936|HumanIA/Henry/1936|USAIH1N1
1935|HumanIA/bh/1935|H1N1
1933|HumanIA/United_Kingdom/1/1933|United_KingdomIH1N1
1935|HumanIA/Melbourne/1935|AustralialH1N1
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1
2004|SwineIA/swine/Zhejiang/1/2004|ChinaIH1N2
2003|SwineIA/swine/Hong_Kong/1763/2003|Hong_KongIH1N2
2002|SwineIA/swine/Hong_Kong/NS13/2002|Hong_KongIH1N2
2000|SwineIA/swine/Hong_Kong/7191/2000|Hong_KongIH1N1
2001|SwineIA/swine/Hong_Kong/8944/2001|Hong_KongIH1N2
1993|SwineIA/swine/Hong_Kong/829/1993|Hong_KongIH1N1
1991|SwineIA/swine/Iowa/24297/1991|USAIH1N1
1998|SwineIA/swine/Wisconsin/464/98|USAIH1N1
1997|SwineIA/swine/Wisconsin/238/97|USAIH1N1
1990|SwineIA/swine/Iowa/3421/1990|USAIH1N1
1988|SwineIA/swine/Wisconsin/1915/1988|USAIH1N1
1987|SwineIA/swine/Iowa/1/1987|USAIH1N1
1986|SwineIA/swine/Iowa/1/1986|USAIH1N1
1980|SwineIA/swine/Wisconsin/1/1980|USAIH1N1
1979|SwineIA/swine/Minnesota/5892-7/1979|USAIH1N1
1977|SwineIA/swine/Tennessee/82/1977|USAIH1N1
1976|SwineIA/swine/Tennessee/15/1976|USAIH1N1
1975|SwineIA/swine/Illinois/1/1975|USAIH1N1
1971|SwineIA/swine/Wisconsin/1/1971|USAIH1N1
1970|SwineIA/swine/Wisconsin/2/1970|USAIH1N1
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1
1931|SwineIA/swine/USA/1976/1931|USAIH1N1
1930|SwineIA/swine/Iowa/15/1930|USAIH1N1
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

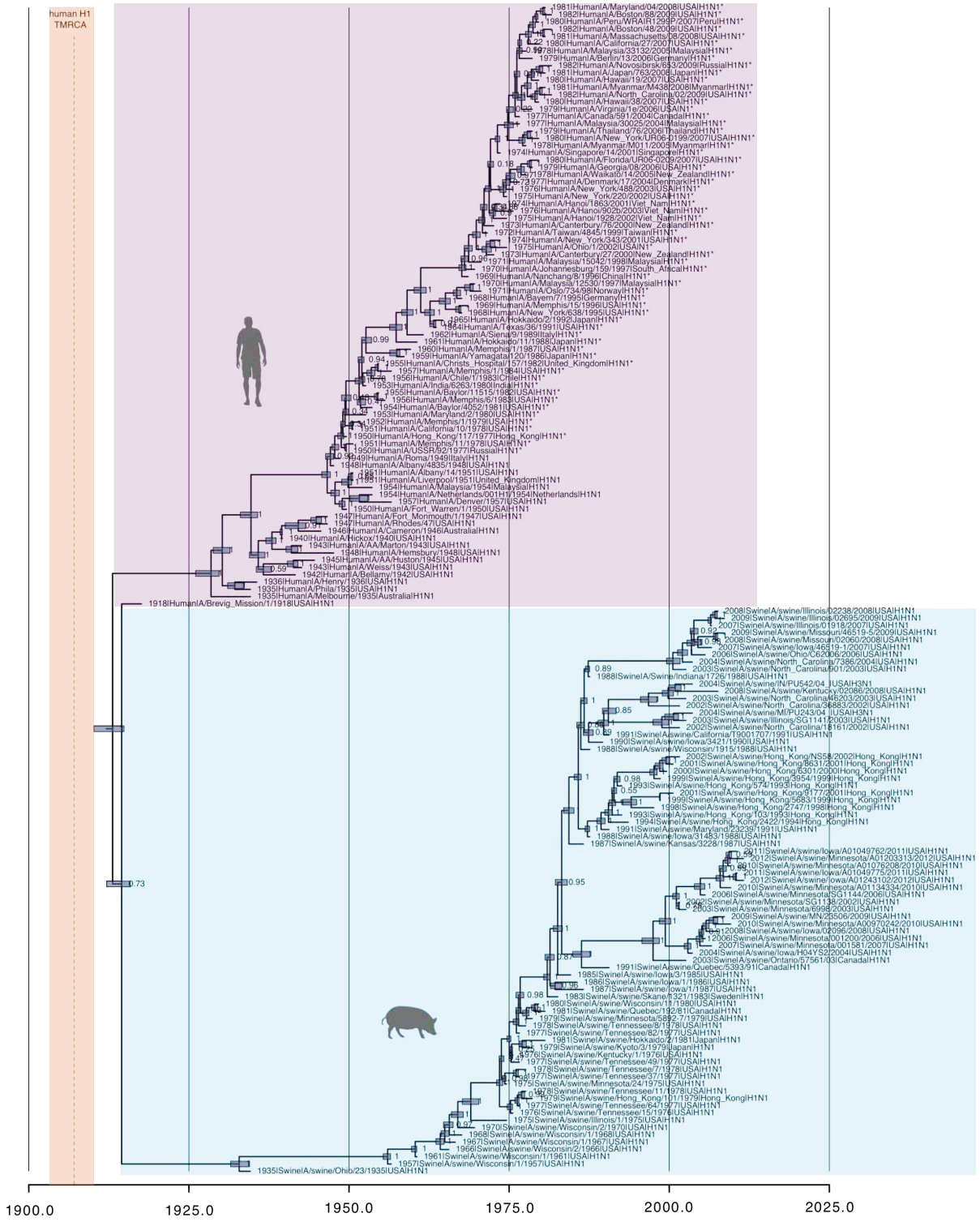
D *HA* (H1)



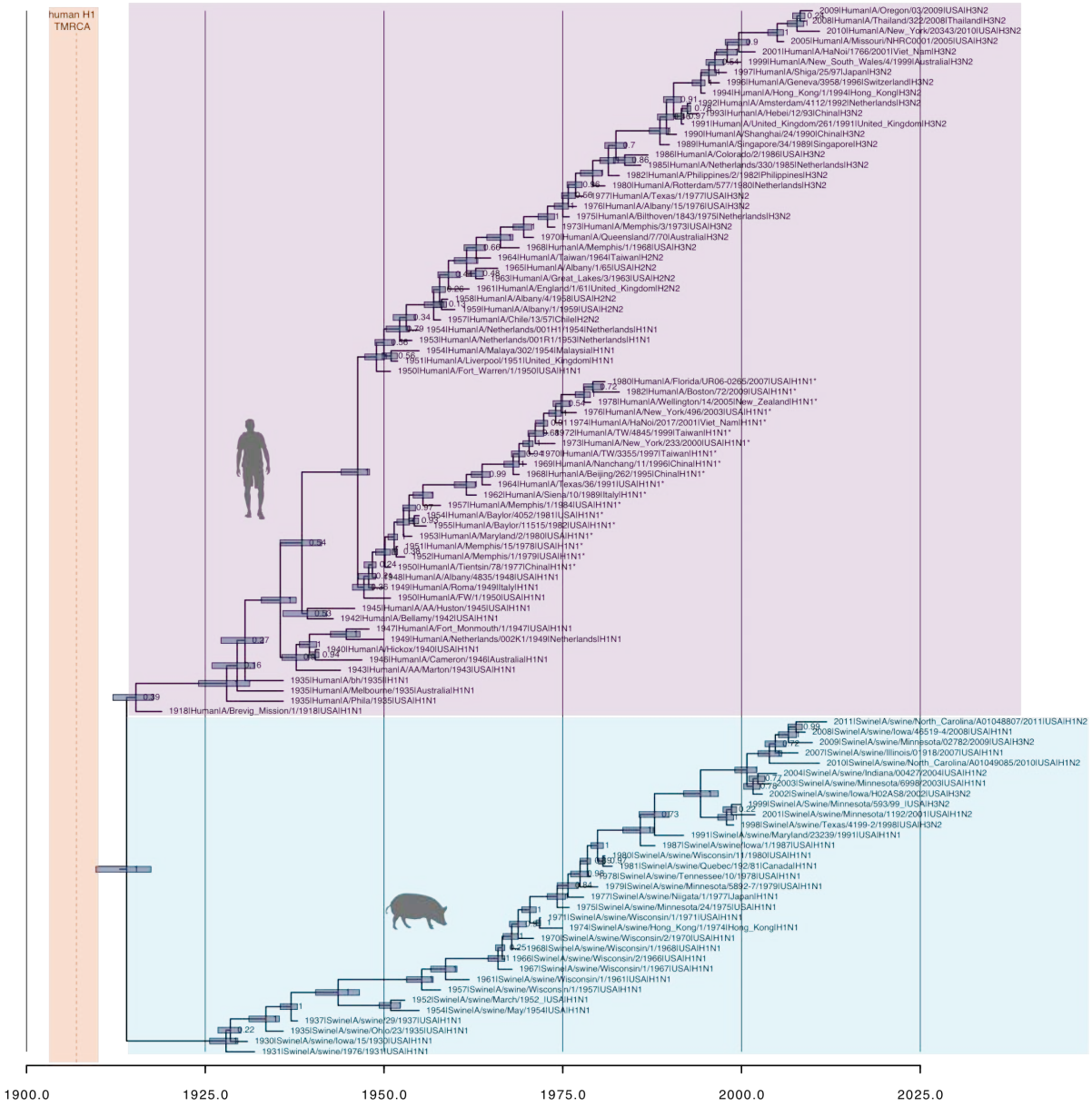
E *NP*


human H1
TMRCA
2009|HumanIA/Nanjing/1/2009|ChinaIH3N2
2011|HumanIA/California/NHRC0002/2011|USAIH3N2
2008|HumanIA/California/UR07-0063/2008|USAIH3N2
2007|HumanIA/Sydney/202G/2007|AustraliaIH3N2
2006|HumanIA/Sydney/186R/2006|AustraliaIH3N2
2005|HumanIA/Texas/NHRC0001/2005|USAIH3N2
2003|HumanIA/Wyoming/03/2003|USAIH3N2
2002|HumanIA/Yokohama/47/2002|JapanIH1N2
1999|HumanIA/Memphis/59/1999|USAIH3N2
1997|HumanIA/Siena/1/1997|ItalyIH3N2
1993|HumanIA/Lyon/23602/1993|FranceIH3N2
1994|HumanIA/New_York/753/1994|USAIH3N2
1989|HumanIA/Eindhoven/3447/1989|NetherlandsIH3N2
1988|HumanIA/Netherlands/450/1988|NetherlandsIH3N2
1984|HumanIA/Hong_Kong/4/1984|Hong_KongIH3N2
1983|HumanIA/Oslo/13676/1983|NorwayIH3N2
1982|HumanIA/Bilthoven/10684/1982|NetherlandsIH3N2
1980|HumanIA/Nanjing/13/1980|ChinaIH3N2
1979|HumanIA/Kiev/59/1979|UkraineIH1N1
1977|HumanIA/Amsterdam/1609/1977|NetherlandsIH3N2
1976|HumanIA/Bilthoven/1761/1976|NetherlandsIH3N2
1975|HumanIA/Hong_Kong/43/1975|Hong_KongIH3N2
1972|HumanIA/Udorn/307/1972|RussiaIH3N2
1973|HumanIA/Memphis/1/1973|USAIH3N2
1971|HumanIA/Bilthoven/21801/1971|NetherlandsIH3N2
1965|HumanIA/Albany/1/65|USAIH2N2
1964|HumanIA/Gottbus/1/1964|GermanyIH2N2
1969|HumanIA/Hong_Kong/3/1969|Hong_KongIH3N2
1970|HumanIA/Albany/1/1970|USAIH3N2
1968|HumanIA/Philippines/104/68|PhilippinesIH3N2
1963|HumanIA/North_Carolina/1/1963|USAIH2N2
1962|HumanIA/Netherlands/60/62|NetherlandsIH2N2
1960|HumanIA/Netherlands/056H1/1960|NetherlandsIH2N2
1961|HumanIA/England/1/61|United_KingdomIH2N2
1958|HumanIA/Albany/6/58|USAIH2N2
1959|HumanIA/Krasnodar/101/1959|RussiaIH2N2
1957|HumanIA/Singapore/1/1957|SingaporeIH2N2
1956|HumanIA/Netherlands/002Z1/1956|NetherlandsIH1N1
1954|HumanIA/Netherlands/001H1/1954|NetherlandsIH1N1
1953|HumanIA/Netherlands/001R1/1953|NetherlandsIH1N1
1951|HumanIA/Albany/13/1951|USAIH1N1
1955|HumanIA/England/19/1955|United_KingdomIH1N1
1954|HumanIA/Malaysia/1954|MalaysiaIH1N1
1981|HumanIA/Florida/UR07-0022/2008|USAIH1N1*
1982|HumanIA/Arkansas/WRAIR1249P/2009|USAIH1N1*
1980|HumanIA/Texas/UR06-0217/2007|USAIH1N1*
1979|HumanIA/New_York/8/2006|USAIH1N1*
1978|HumanIA/Denmark/22/2005|DenmarkIH1N1*
1973|HumanIA/South_Canterbury/50/2000|New_ZealandIH1N1*
1972|HumanIA/New_South_Wales/24/1999|AustraliaIH1N1*
1969|HumanIA/New_York/626/1996|USAIH1N1*
1968|HumanIA/Siena/14/1995|ItalyIH1N1*
1964|HumanIA/Texas/36/1991|USAIH1N1*
1962|HumanIA/Siena/10/1989|ItalyIH1N1*
1959|HumanIA/Singapore/6/1986|SingaporeIH1N1*
1960|HumanIA/Siena/4/1987|ItalyIH1N1*
1954|HumanIA/Baylor/4052/1981|USAIH1N1*
1955|HumanIA/Baylor/11515/1982|USAIH1N1*
1956|HumanIA/Memphis/49/1985|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1951|HumanIA/Memphis/10/1978|USAIH1N1*
1950|HumanIA/Tientsin/78/1977|ChinaIH1N1*
1950|HumanIA/Fort_Warren/1/1950|USAIH1N1
1949|HumanIA/Roma/1949|ItalyIH1N1
1948|HumanIA/Albany/4835/1948|USAIH1N1
1947|HumanIA/Monmouth/1947|USAIH1N1
1949|HumanIA/Netherlands/002K1/1949|NetherlandsIH1N1
1940|HumanIA/Hickox/1940|USAIH1N1
1945|HumanIA/Cameron/1946|AustraliaIH1N1
1943|HumanIA/Weiss/1943|USAIH1N1
1945|HumanIA/AA/Huston/1945|USAIH1N1
1942|HumanIA/Bellamy/1942|USAIH1N1
1948|HumanIA/Hemsbury/1948|USAIH1N1
1943|HumanIA/AA/Marton/1943|USAIH1N1
1936|HumanIA/Henry/1936|USAIH1N1
1935|HumanIA/Phila/1935|USAIH1N1
1935|HumanIA/Bth/1935|IH1N1
1935|HumanIA/Melbourne/1935|AustraliaIH1N1
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1
2008|SwineIA/swine/Illinois/01981/2008|USAIH1N1
2009|SwineIA/swine/Illinois/02510/2009|USAIH1N1
2010|SwineIA/swine/Pennsylvania/057108-1/2010|USAIH3N2
2001|SwineIA/Swine/North_Carolina/93523/01_IUSAIH1N2
2000|SwineIA/swine/North_Carolina/47834/2000|USAIH1N1
2005|SwineIA/swine/Guangxi/17/2005|ChinaIH1N2
2006|SwineIA/swine/Guangxi/132/2006|ChinaIH1N2
2003|SwineIA/swine/Illinois/SG1141/2003|USAIH1N1
2002|SwineIA/swine/North_Carolina/18161/2002|USAIH1N1
2007|SwineIA/swine/IL/07003243/2007|USAIH1N2
2008|SwineIA/swine/Texas/008648/2008|USAIH1N2
2004|SwineIA/swine/Korea/CAS05/2004|South_KoreaIH3N2
2001|SwineIA/swine/Minnesota/1192/2001|USAIH1N2
2000|SwineIA/Swine/Minnesota/55551/2000|USAIH1N2
1999|SwineIA/Swine/Iowa/533/99_IUSAIH3N2
1990|SwineIA/swine/Iowa/3421/1990|USAIH1N1
2002|SwineIA/swine/Hong_Kong/NS728/2002|Hong_KongIH1N2
2003|SwineIA/swine/Hong_Kong/NS605/2003|Hong_KongIH1N2
2004|SwineIA/swine/Zhejiang/1/2004|ChinaIH1N2
2001|SwineIA/swine/Hong_Kong/8944/2001|Hong_KongIH1N2
1999|SwineIA/swine/Hong_Kong/5609/1999|Hong_KongIH1N2
1993|SwineIA/swine/Hong_Kong/103/1993|Hong_KongIH1N1
1994|SwineIA/swine/Hong_Kong/1795/1994|Hong_KongIH1N1
1991|SwineIA/swine/Maryland/23239/1991|USAIH1N1
1988|SwineIA/swine/Iowa/17672/1988|USAIH1N1
1983|SwineIA/swine/Skane/1321/1983|SwedenIH1N1
1986|SwineIA/swine/Iowa/1/1986|USAIH1N1
1980|SwineIA/swine/Wisconsin/1/1980|USAIH1N1
1981|SwineIA/swine/Quebec/192/81|CanadaIH1N1
1979|SwineIA/swine/Hong_Kong/103/1979|Hong_KongIH1N1
1976|SwineIA/swine/Iowa/1/1976|USAIH1N1
1979|SwineIA/swine/Kyoto/3/1979|JapanIH1N1
1981|SwineIA/swine/Hokkaido/2/1981|JapanIH1N1
1980|SwineIA/swine/Ehime/1/1980|JapanIH1N2
1977|SwineIA/swine/Arizona/148/1977|USAIH1N1
1977|SwineIA/swine/Nebraska/123/1977|USAIH1N1
1971|SwineIA/swine/Wisconsin/1/1971|USAIH1N1
1975|SwineIA/swine/Illinois/1/1975|USAIH1N1
1966|SwineIA/swine/Wisconsin/2/1966|USAIH1N1
1970|SwineIA/swine/Wisconsin/2/1970|USAIH1N1
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1
1937|SwineIA/swine/29/1937|USAIH1N1
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1
1931|SwineIA/swine/USA/1976/1931|USAIH1N1
1930|SwineIA/swine/Iowa/15/1930|USAIH1N1
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

F *NA* (N1)


human H1
TMRCA
1981|HumanIA/Maryland/04/2008|USAIH1N1*
1982|HumanIA/Boston/88/2009|USAIH1N1*
1980|HumanIA/Peru/WRAIR1299P/2007|PeruIH1N1*
1982|HumanIA/Boston/48/2009|USAIH1N1*
1981|HumanIA/Massachusetts/08/2008|USAIH1N1*
1980|HumanIA/California/27/2007|USAIH1N1*
1978|HumanIA/Malaysia/33132/2005|MalaysiaIH1N1*
1979|HumanIA/Berlin/13/2006|GermanyIH1N1*
1982|HumanIA/Novosibirsk/693/2009|RussiaIH1N1*
1981|HumanIA/Japan/763/2008|JapanIH1N1*
1980|HumanIA/Hawaii/19/2007|USAIH1N1*
1981|HumanIA/Myanmar/M438/2008|MyanmarIH1N1*
1982|HumanIA/North_Carolina/02/2009|USAIH1N1*
1980|HumanIA/Hawaii/38/2007|USAIH1N1*
1979|HumanIA/Virginia/1e/2006|USAIN1*
1977|HumanIA/Canada/591/2004|CanadaIH1N1*
1977|HumanIA/Malaysia/30025/2004|MalaysiaIH1N1*
1979|HumanIA/Thailand/76/2006|ThailandIH1N1*
1980|HumanIA/New_York/UR06-0199/2007|USAIH1N1*
1978|HumanIA/Myanmar/M011/2005|MyanmarIH1N1*
1974|HumanIA/Singapore/14/2001|SingaporeIH1N1*
1980|HumanIA/Florida/UR06-0209/2007|USAIH1N1*
1979|HumanIA/Georgia/08/2006|USAIH1N1*
1978|HumanIA/Waikato/5/14/2005|New_ZealandIH1N1*
1977|HumanIA/Denmark/17/2004|DenmarkIH1N1*
1976|HumanIA/New_York/488/2003|USAIH1N1*
1975|HumanIA/New_York/220/2002|USAIH1N1*
1974|HumanIA/Hanoi/1863/2001|Viet_NamIH1N1*
1976|HumanIA/Hanoi/902b/2003|Viet_NamIH1N1*
1975|HumanIA/Hanoi/1928/2002|Viet_NamIH1N1*
1973|HumanIA/Canterbury/76/2000|New_ZealandIH1N1*
1972|HumanIA/Taiwan/4845/1999|TaiwanIH1N1*
1974|HumanIA/New_York/343/2001|USAIH1N1*
1975|HumanIA/Ohio/1/2002|USAIN1*
1973|HumanIA/Canterbury/27/2000|New_ZealandIH1N1*
1971|HumanIA/Malaysia/1504/1998|MalaysiaIH1N1*
1970|HumanIA/Johannesburg/159/1997|South_AfricaIH1N1*
1969|HumanIA/Nanchang/8/1996|ChinaIH1N1*
1970|HumanIA/Malaysia/12530/1997|MalaysiaIH1N1*
1971|HumanIA/Oslo/734/98|NorwayIH1N1*
1968|HumanIA/Bayern/7/1995|GermanyIH1N1*
1969|HumanIA/Memphis/15/1996|USAIH1N1*
1968|HumanIA/New_York/638/1995|USAIH1N1*
1965|HumanIA/Hokkaido/2/1992|JapanIH1N1*
1964|HumanIA/Texas/36/1991|USAIH1N1*
1962|HumanIA/Siena/9/1989|ItalyIH1N1*
1961|HumanIA/Hokkaido/11/1988|JapanIH1N1*
1960|HumanIA/Memphis/1/1987|USAIH1N1*
1959|HumanIA/Yamagata/120/1986|JapanIH1N1*
1955|HumanIA/Christs_Hospital/157/1982|United_KingdomIH1N1*
1957|HumanIA/Memphis/1/1984|USAIH1N1*
1956|HumanIA/Chile/1/1983|ChileIH1N1*
1953|HumanIA/India/6263/1980|IndiaIH1N1*
1955|HumanIA/Baylor/11515/1982|USAIH1N1*
1956|HumanIA/Memphis/6/1983|USAIH1N1*
1954|HumanIA/Baylor/4052/1981|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1952|HumanIA/Memphis/1/1979|USAIH1N1*
1951|HumanIA/California/10/1978|USAIH1N1*
1950|HumanIA/Hong_Kong/117/1977|Hong_KongIH1N1*
1951|HumanIA/Memphis/11/1978|USAIH1N1*
1950|HumanIA/USSR/92/1977|RussiaIH1N1*
1949|HumanIA/Roma/1949|ItalyIH1N1*
1948|HumanIA/Albany/4835/1948|USAIH1N1*
1951|HumanIA/Albany/14/1951|USAIH1N1*
1951|HumanIA/Liverpool/1951|United_KingdomIH1N1*
1954|HumanIA/Malaysia/1954|MalaysiaIH1N1*
1954|HumanIA/Netherlands/001/1954|NetherlandsIH1N1*
1957|HumanIA/Denver/1957|USAIH1N1*
1950|HumanIA/Fort_Warren/1/1950|USAIH1N1*
1947|HumanIA/Fort_Monmouth/1/1947|USAIH1N1*
1947|HumanIA/Rhodes/47|USAIH1N1*
1946|HumanIA/Cameron/1946|AustraliaIH1N1*
1940|HumanIA/Hickox/1940|USAIH1N1*
1943|HumanIA/AA/Marton/1943|USAIH1N1*
1948|HumanIA/Hemsbury/1948|USAIH1N1*
1945|HumanIA/AA/Huston/1945|USAIH1N1*
1943|HumanIA/Weiss/1943|USAIH1N1*
1942|HumanIA/Bellamy/1942|USAIH1N1*
1936|HumanIA/Henry/1936|USAIH1N1*
1935|HumanIA/Phila/1935|USAIH1N1*
1935|HumanIA/Melbourne/1935|AustraliaIH1N1*
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1*
2008|SwineIA/swine/Illinois/02238/2008|USAIH1N1*
2009|SwineIA/swine/Illinois/02695/2009|USAIH1N1*
2007|SwineIA/swine/Illinois/01918/2007|USAIH1N1*
2009|SwineIA/swine/Missouri/46519-5/2009|USAIH1N1*
2008|SwineIA/swine/Missouri/02060/2008|USAIH1N1*
2007|SwineIA/swine/Iowa/46519-1/2007|USAIH1N1*
2006|SwineIA/swine/Ohio/C62006/2006|USAIH1N1*
2004|SwineIA/swine/North_Carolina/7386/2004|USAIH1N1*
2003|SwineIA/swine/North_Carolina/601/2003|USAIH1N1*
1988|SwineIA/Swine/Indiana/1726/1988|USAIH1N1*
2004|SwineIA/swine/IN/PU542/04_1|USAIH3N1*
2008|SwineIA/swine/Kentucky/02086/2008|USAIH1N1*
2003|SwineIA/swine/North_Carolina/46203/2003|USAIH1N1*
2002|SwineIA/swine/North_Carolina/36883/2002|USAIH1N1*
2004|SwineIA/swine/MI/PU243/04_1|USAIH3N1*
2003|SwineIA/swine/Illinois/SG11417/2003|USAIH1N1*
2002|SwineIA/swine/North_Carolina/18161/2002|USAIH1N1*
1991|SwineIA/swine/California/T90017071991|USAIH1N1*
1990|SwineIA/swine/Iowa/3421/1990|USAIH1N1*
1988|SwineIA/swine/Wisconsin/1915/1988|USAIH1N1*
2002|SwineIA/swine/Hong_Kong/NS58/2002|Hong_KongIH1N1*
2001|SwineIA/swine/Hong_Kong/8631/2001|Hong_KongIH1N1*
2000|SwineIA/swine/Hong_Kong/6301/2000|Hong_KongIH1N1*
1999|SwineIA/swine/Hong_Kong/3954/1999|Hong_KongIH1N1*
1993|SwineIA/swine/Hong_Kong/574/1993|Hong_KongIH1N1*
2001|SwineIA/swine/Hong_Kong/9177/2001|Hong_KongIH1N1*
1999|SwineIA/swine/Hong_Kong/5683/1999|Hong_KongIH1N1*
1998|SwineIA/swine/Hong_Kong/2747/1998|Hong_KongIH1N1*
1993|SwineIA/swine/Hong_Kong/103/1993|Hong_KongIH1N1*
1994|SwineIA/swine/Hong_Kong/2422/1994|Hong_KongIH1N1*
1991|SwineIA/swine/Maryland/23239/1991|USAIH1N1*
1988|SwineIA/swine/Iowa/31453/1988|USAIH1N1*
1987|SwineIA/swine/Kansas/3228/1987|USAIH1N1*
2011|SwineIA/swine/Iowa/A01049762/2011|USAIH1N1*
2012|SwineIA/swine/Minnesota/A01203313/2012|USAIH1N1*
2010|SwineIA/swine/Minnesota/A01076208/2010|USAIH1N1*
2011|SwineIA/swine/Iowa/A01049775/2011|USAIH1N1*
2012|SwineIA/swine/Iowa/A01243102/2012|USAIH1N1*
2010|SwineIA/swine/Minnesota/A01134334/2010|USAIH1N1*
2006|SwineIA/swine/Minnesota/SG1144/2006|USAIH1N1*
2002|SwineIA/swine/Minnesota/SG1138/2002|USAIH1N1*
2003|SwineIA/swine/Minnesota/6998/2003|USAIH1N1*
2009|SwineIA/swine/MN/29506/2009|USAIH1N1*
2010|SwineIA/swine/Minnesota/A00970742/2010|USAIH1N1*
2008|SwineIA/swine/Iowa/02096/2008|USAIH1N1*
2006|SwineIA/swine/Minnesota/001200/2006|USAIH1N1*
2007|SwineIA/swine/Minnesota/001581/2007|USAIH1N1*
2004|SwineIA/swine/Iowa/H04YS2|2004|USAIH1N1*
2003|SwineIA/swine/Ontario/57561/03|CanadaIH1N1*
1991|SwineIA/swine/Quebec/5393/91|CanadaIH1N1*
1985|SwineIA/swine/Iowa/3/1985|USAIH1N1*
1986|SwineIA/swine/Iowa/1/1986|USAIH1N1*
1987|SwineIA/swine/Iowa/1/1987|USAIH1N1*
1983|SwineIA/swine/Skane/1321/1983|SwedenIH1N1*
1980|SwineIA/swine/Wisconsin/1/1980|USAIH1N1*
1981|SwineIA/swine/Quebec/192/81|CanadaIH1N1*
1979|SwineIA/swine/Minnesota/5892-7/1979|USAIH1N1*
1978|SwineIA/swine/Tennessee/8/1978|USAIH1N1*
1977|SwineIA/swine/Tennessee/82/1977|USAIH1N1*
1981|SwineIA/swine/Hokkaido/2/1981|JapanIH1N1*
1979|SwineIA/swine/Kyoto/3/1979|JapanIH1N1*
1976|SwineIA/swine/Kentucky/1/1976|USAIH1N1*
1977|SwineIA/swine/Tennessee/49/1977|USAIH1N1*
1978|SwineIA/swine/Tennessee/7/1978|USAIH1N1*
1977|SwineIA/swine/Tennessee/37/1977|USAIH1N1*
1975|SwineIA/swine/Minnesota/24/1975|USAIH1N1*
1978|SwineIA/swine/Tennessee/11/1978|USAIH1N1*
1979|SwineIA/swine/Hong_Kong/101/1979|Hong_KongIH1N1*
1977|SwineIA/swine/Tennessee/64/1977|USAIH1N1*
1976|SwineIA/swine/Tennessee/15/1976|USAIH1N1*
1975|SwineIA/swine/Illinois/1/1975|USAIH1N1*
1970|SwineIA/swine/Wisconsin/2/1970|USAIH1N1*
1968|SwineIA/swine/Wisconsin/1/1968|USAIH1N1*
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1*
1966|SwineIA/swine/Wisconsin/2/1966|USAIH1N1*
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1*
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1*
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1*
1.00
0.99
0.94
0.96
0.89
0.85
0.73
0.95
0.98
0.87
0.97
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

G *M1/2*


human H1
TMRCA
2009|HumanIA/Oregon/03/2009|USAIH3N2
2008|HumanIA/Thailand/322/2008|ThailandIH3N2
2010|HumanIA/New_York/20343/2010|USAIH3N2
2005|HumanIA/Missouri/NHRC0001/2005|USAIH3N2
2001|HumanIA/HaNoi/1766/2001|Viet_NamIH3N2
1999|HumanIA/New_South_Wales/4/1999|AustraliaIH3N2
1997|HumanIA/Shiga/25/97|JapanIH3N2
1996|HumanIA/Geneva/3958/1996|SwitzerlandIH3N2
1994|HumanIA/Hong_Kong/1/1994|Hong_KongIH3N2
1992|HumanIA/Amsterdam/4112/1992|NetherlandsIH3N2
1991|HumanIA/United_Kingdom/261/1991|United_KingdomIH3N2
1990|HumanIA/Shanghai/24/1990|ChinaIH3N2
1989|HumanIA/Singapore/34/1989|SingaporeIH3N2
1986|HumanIA/Colorado/2/1986|USAIH3N2
1985|HumanIA/Netherlands/330/1985|NetherlandsIH3N2
1982|HumanIA/Philippines/2/1982|PhilippinesIH3N2
1980|HumanIA/Rotterdam/577/1980|NetherlandsIH3N2
1977|HumanIA/Texas/1/1977|USAIH3N2
1976|HumanIA/Albany/15/1976|USAIH3N2
1975|HumanIA/Bilthoven/1843/1975|NetherlandsIH3N2
1973|HumanIA/Memphis/3/1973|USAIH3N2
1970|HumanIA/Queensland/7/70|AustraliaIH3N2
1968|HumanIA/Memphis/1/1968|USAIH3N2
1964|HumanIA/Taiwan/1964|TaiwanIH2N2
1965|HumanIA/Albany/1/65|USAIH2N2
1963|HumanIA/Great_Lakes/3/1963|USAIH2N2
1961|HumanIA/England/1/61|United_KingdomIH2N2
1958|HumanIA/Albany/4/1958|USAIH2N2
1959|HumanIA/Albany/1/1959|USAIH2N2
1957|HumanIA/Chile/13/57|ChileIH2N2
1954|HumanIA/Netherlands/001H1|1954|NetherlandsIH1N1
1953|HumanIA/Netherlands/001R1/1953|NetherlandsIH1N1
1954|HumanIA/Malaya/302/1954|MalaysiaIH1N1
1951|HumanIA/Liverpool/1951|United_KingdomIH1N1
1950|HumanIA/Fort_Warren/1/1950|USAIH1N1
1980|HumanIA/Florida/UR06-0265/2007|USAIH1N1*
1982|HumanIA/Boston/72/2009|USAIH1N1*
1978|HumanIA/Wellington/14/2005|New_ZealandIH1N1*
1976|HumanIA/New_York/496/2003|USAIH1N1*
1974|HumanIA/HaNoi/2017/2001|Viet_NamIH1N1*
1972|HumanIA/TW/4845/1999|TaiwanIH1N1*
1973|HumanIA/New_York/233/2000|USAIH1N1*
1970|HumanIA/TW/3355/1997|TaiwanIH1N1*
1969|HumanIA/Nanchang/11/1996|ChinaIH1N1*
1968|HumanIA/Beijing/262/1995|ChinaIH1N1*
1964|HumanIA/Texas/36/1991|USAIH1N1*
1962|HumanIA/Siena/10/1989|ItalyIH1N1*
1957|HumanIA/Memphis/1/1984|USAIH1N1*
1954|HumanIA/Baylor/4052/1981|USAIH1N1*
1955|HumanIA/Baylor/11515/1982|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1951|HumanIA/Memphis/15/1978|USAIH1N1*
1952|HumanIA/Memphis/1/1979|USAIH1N1*
1950|HumanIA/Tientsin/78/1977|ChinaIH1N1*
1948|HumanIA/Albany/4835/1948|USAIH1N1
1949|HumanIA/Roma/1949|ItalyIH1N1
1950|HumanIA/FW/1/1950|USAIH1N1
1945|HumanIA/AA/Huston/1945|USAIH1N1
1942|HumanIA/Bellamy/1942|USAIH1N1
1947|HumanIA/Fort_Monmouth/1/1947|USAIH1N1
1949|HumanIA/Netherlands/002K1/1949|NetherlandsIH1N1
1940|HumanIA/Hickox/1940|USAIH1N1
1946|HumanIA/Cameron/1946|AustraliaIH1N1
1943|HumanIA/AA/Marton/1943|USAIH1N1
1935|HumanIA/bh/1935|IH1N1
1935|HumanIA/Melbourne/1935|AustraliaIH1N1
1935|HumanIA/Phila/1935|USAIH1N1
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1
2011|SwineIA/swine/North_Carolina/A01048807/2011|USAIH1N2
2008|SwineIA/swine/Iowa/46519-4/2008|USAIH1N1
2009|SwineIA/swine/Minnesota/02782/2009|USAIH3N2
2007|SwineIA/swine/Illinois/01918/2007|USAIH1N1
2010|SwineIA/swine/North_Carolina/A01049085/2010|USAIH1N2
2004|SwineIA/swine/Indiana/00427/2004|USAIH1N2
2003|SwineIA/swine/Minnesota/6998/2003|USAIH1N1
2002|SwineIA/swine/Iowa/H02AS8/2002|USAIH3N2
1999|SwineIA/Swine/Minnesota/593/99_IUSAIH3N2
2001|SwineIA/swine/Minnesota/1192/2001|USAIH1N2
1998|SwineIA/swine/Texas/4199-2/1998|USAIH3N2
1991|SwineIA/swine/Maryland/23239/1991|USAIH1N1
1987|SwineIA/swine/Iowa/1/1987|USAIH1N1
1980|SwineIA/swine/Wisconsin/1/1980|USAIH1N1
1981|SwineIA/swine/Quebec/192/81|CanadaIH1N1
1978|SwineIA/swine/Tennessee/10/1978|USAIH1N1
1979|SwineIA/swine/Minnesota/5892-7/1979|USAIH1N1
1977|SwineIA/swine/Niigata/1/1977|JapanIH1N1
1975|SwineIA/swine/Minnesota/24/1975|USAIH1N1
1971|SwineIA/swine/Wisconsin/1/1971|USAIH1N1
1974|SwineIA/swine/Hong_Kong/1/1974|Hong_KongIH1N1
1970|SwineIA/swine/Wisconsin/2/1970|USAIH1N1
1968|SwineIA/swine/Wisconsin/1/1968|USAIH1N1
1966|SwineIA/swine/Wisconsin/2/1966|USAIH1N1
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1
1952|SwineIA/swine/March/1952_IUSAIH1N1
1954|SwineIA/swine/May/1954|USAIH1N1
1937|SwineIA/swine/29/1937|USAIH1N1
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1
1930|SwineIA/swine/Iowa/15/1930|USAIH1N1
1931|SwineIA/swine/1976/1931|USAIH1N1
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

H *NS1/2*

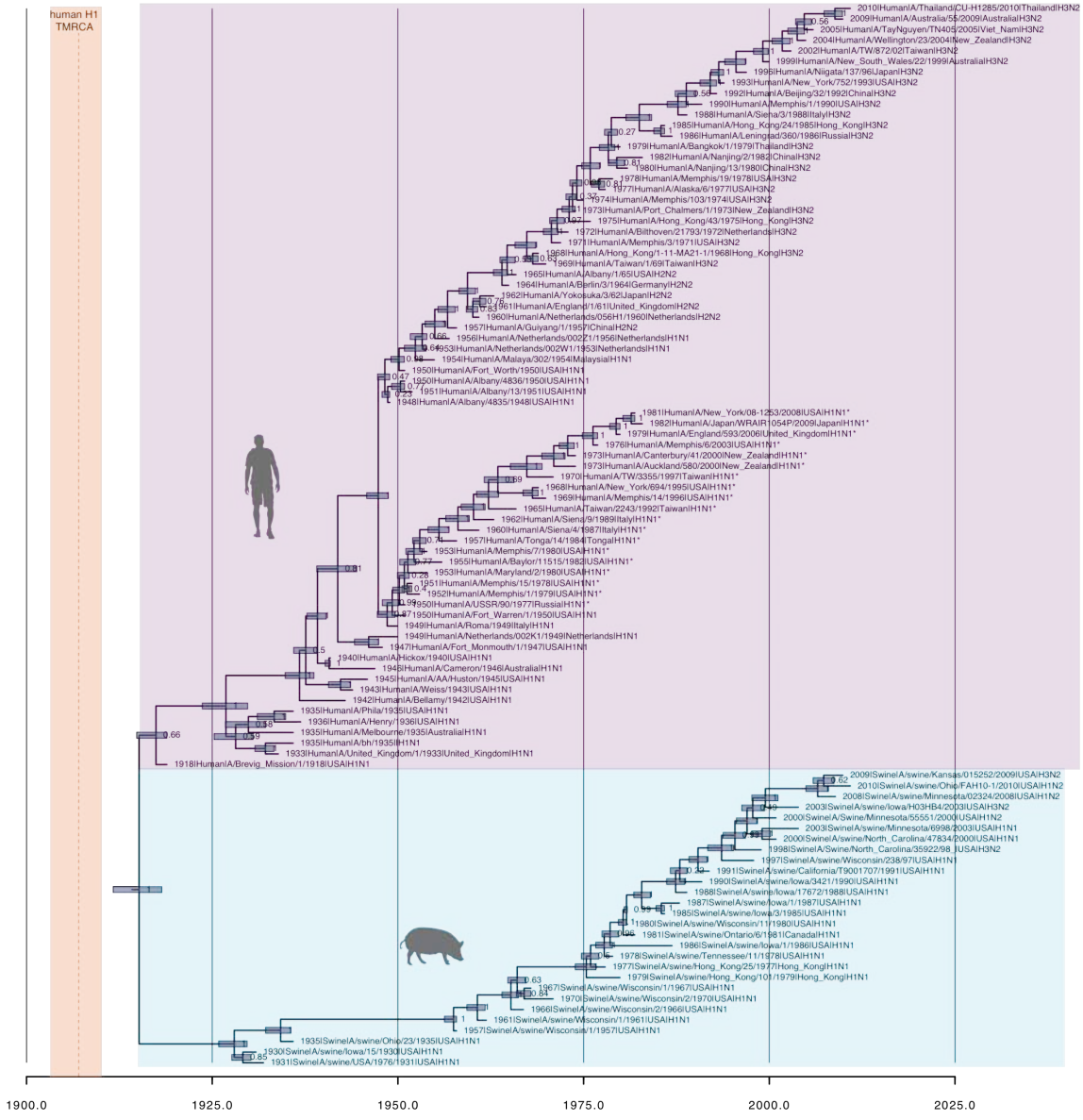

human H1 TMRCA
2010|HumanIA/Thailand/CU-H1285/2010|ThailandIH3N2
2009|HumanIA/Australia/55/2009|AustraliaIH3N2
2005|HumanIA/TayNguyen/TN405/2005|Viet_NamIH3N2
2004|HumanIA/Wellington/23/2004|New_ZealandIH3N2
2002|HumanIA/TW/872/02|TaiwanIH3N2
1999|HumanIA/New_South_Wales/22/1999|AustraliaIH3N2
1998|HumanIA/Niigata/137/96|JapanIH3N2
1993|HumanIA/New_York/752/1993|USAIH3N2
1992|HumanIA/Beijing/32/1992|ChinaIH3N2
1990|HumanIA/Memphis/1/1990|USAIH3N2
1988|HumanIA/Sichu/3/1988|ItalyIH3N2
1985|HumanIA/Hong_Kong/24/1985|Hong_KongIH3N2
1986|HumanIA/Leningrad/360/1986|RussiaIH3N2
1979|HumanIA/Bangkok/1/1979|ThailandIH3N2
1982|HumanIA/Nanjing/2/1982|ChinaIH3N2
1980|HumanIA/Nanjing/13/1980|ChinaIH3N2
1978|HumanIA/Memphis/19/1978|USAIH3N2
1977|HumanIA/Alaska/6/1977|USAIH3N2
1974|HumanIA/Memphis/103/1974|USAIH3N2
1973|HumanIA/Port_Chalmers/1/1973|New_ZealandH3N2
1975|HumanIA/Hong_Kong/43/1975|Hong_KongIH3N2
1972|HumanIA/Bilthoven/21793/1972|NetherlandsIH3N2
1971|HumanIA/Memphis/3/1971|USAIH3N2
1968|HumanIA/Hong_Kong/1-11-MA21-1/1968|Hong_KongIH3N2
1969|HumanIA/Taiwan/1/69|TaiwanIH3N2
1965|HumanIA/Albany/1/65|USAIH2N2
1964|HumanIA/Berlin/3/1964|GermanyIH2N2
1962|HumanIA/Yokosuka/3/62|JapanIH2N2
1961|HumanIA/England/1/61|United_KingdomIH2N2
1960|HumanIA/Netherlands/056H1/1960|NetherlandsIH2N2
1957|HumanIA/Guiyang/1/1957|ChinaIH2N2
1956|HumanIA/Netherlands/002Z1/1956|NetherlandsIH1N1
1953|HumanIA/Netherlands/002W1/1953|NetherlandsIH1N1
1954|HumanIA/Malaya/302/1954|MalaysiaIH1N1
1950|HumanIA/Fort_Worth/1950|USAIH1N1
1950|HumanIA/Albany/4836/1950|USAIH1N1
1951|HumanIA/Albany/13/1951|USAIH1N1
1948|HumanIA/Albany/4835/1948|USAIH1N1
1981|HumanIA/New_York/08-1253/2008|USAIH1N1*
1982|HumanIA/Japan/WRAIR1054P/2009|JapanIH1N1*
1979|HumanIA/England/593/2006|United_KingdomIH1N1*
1976|HumanIA/Memphis/6/2003|USAIH1N1*
1973|HumanIA/Canterbury/41/2000|New_ZealandIH1N1*
1973|HumanIA/Auckland/580/2000|New_ZealandIH1N1*
1970|HumanIA/TW/3355/1997|TaiwanIH1N1*
1968|HumanIA/New_York/694/1995|USAIH1N1*
1969|HumanIA/Memphis/14/1996|USAIH1N1*
1965|HumanIA/Taiwan/2243/1992|TaiwanIH1N1*
1962|HumanIA/Siena/9/1989|ItalyIH1N1*
1960|HumanIA/Siena/4/1987|ItalyIH1N1*
1957|HumanIA/Tonga/14/1984|TongaIH1N1*
1953|HumanIA/Memphis/7/1980|USAIH1N1*
1955|HumanIA/Baylor/11515/1982|USAIH1N1*
1953|HumanIA/Maryland/2/1980|USAIH1N1*
1951|HumanIA/Memphis/15/1978|USAIH1N1*
1952|HumanIA/Memphis/1/1979|USAIH1N1*
1950|HumanIA/USSR/90/1977|RussiaIH1N1*
1950|HumanIA/Fort_Warren/1/1950|USAIH1N1
1949|HumanIA/Roma/1949|ItalyIH1N1
1949|HumanIA/Netherlands/002K1/1949|NetherlandsIH1N1
1947|HumanIA/Fort_Monmouth/1/1947|USAIH1N1
1940|HumanIA/Hickox/1940|USAIH1N1
1943|HumanIA/Cameron/1946|AustraliaIH1N1
1945|HumanIA/AA/Huston/1945|USAIH1N1
1943|HumanIA/Weiss/1943|USAIH1N1
1942|HumanIA/Bellamy/1942|USAIH1N1
1935|HumanIA/Phila/1935|USAIH1N1
1936|HumanIA/Henry/1936|USAIH1N1
1935|HumanIA/Melbourne/1935|AustraliaIH1N1
1935|HumanIA/bh/1935|IH1N1
1933|HumanIA/United_Kingdom/1/1933|United_KingdomIH1N1
1918|HumanIA/Brevig_Mission/1/1918|USAIH1N1
2009|SwineIA/swine/Kansas/015252/2009|USAIH3N2
2010|SwineIA/swine/Ohio/FAH10-1/2010|USAIH1N2
2008|SwineIA/swine/Minnesota/02324/2008|USAIH1N2
2003|SwineIA/swine/Iowa/H03HB4/2003|USAIH3N2
2000|SwineIA/Swine/Minnesota/55551/2000|USAIH1N2
2003|SwineIA/swine/Minnesota/6998/2003|USAIH1N1
2000|SwineIA/swine/North_Carolina/47834/2000|USAIH1N1
1998|SwineIA/Swine/North_Carolina/35922/98_|USAIH3N2
1997|SwineIA/swine/Wisconsin/238/97|USAIH1N1
1991|SwineIA/swine/California/T9001707/1991|USAIH1N1
1990|SwineIA/swine/Iowa/3421/1990|USAIH1N1
1988|SwineIA/swine/Iowa/17672/1988|USAIH1N1
1987|SwineIA/swine/Iowa/1/1987|USAIH1N1
1985|SwineIA/swine/Iowa/3/1985|USAIH1N1
1980|SwineIA/swine/Wisconsin/11/1980|USAIH1N1
1981|SwineIA/swine/Ontario/6/1981|CanadaIH1N1
1986|SwineIA/swine/Iowa/1/1986|USAIH1N1
1978|SwineIA/swine/Tennessee/11/1978|USAIH1N1
1977|SwineIA/swine/Hong_Kong/25/1977|Hong_KongIH1N1
1979|SwineIA/swine/Hong_Kong/101/1979|Hong_KongIH1N1
1967|SwineIA/swine/Wisconsin/1/1967|USAIH1N1
1970|SwineIA/swine/Wisconsin/2/1970|USAIH1N1
1966|SwineIA/swine/Wisconsin/2/1966|USAIH1N1
1961|SwineIA/swine/Wisconsin/1/1961|USAIH1N1
1957|SwineIA/swine/Wisconsin/1/1957|USAIH1N1
1935|SwineIA/swine/Ohio/23/1935|USAIH1N1
1930|SwineIA/swine/Iowa/15/1930|USAIH1N1
1931|SwineIA/swine/USA/1976/1931|USAIH1N1
1900.0
1925.0
1950.0
1975.0
2000.0
2025.0

Fig. S12. Schematic diagrams of various host-jumping hypotheses leading to observed H1 HA variation.

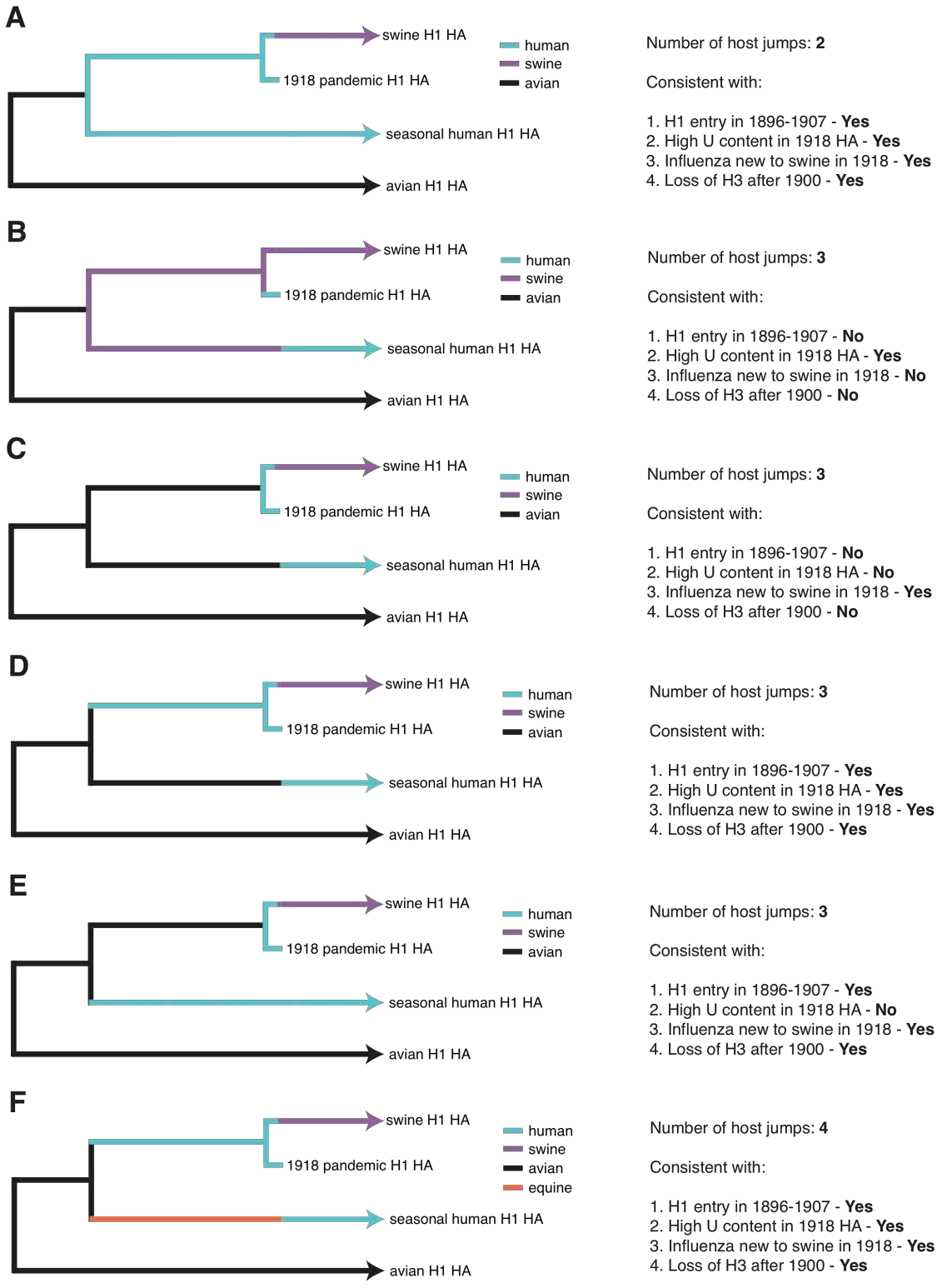


Fig. S13. Maximum clade credibility (MCC) tree of the H1, H2, and H5 subtypes of HA. To the right of the MCC tree are the clade-specific rate distributions (in substitutions/site/year) inferred under the local clock model. The pandemics of 1830-33, 1847-51, and 1889-1893 are shown with gray bars. The posterior probability of each node and 95% CIs on node times are shown. The orange stars indicate points at which H1 or H1-like viruses are assumed to have emerged in humans (in 1830, ~1900, and 2009; the 1977 re-emergence is not marked). The cross represents the putative extinction of the 1830-1889 H1-like HA. That lineage, and the classical swine influenza lineage, represented by dashed branches, are superimposed on the tree for purposes of illustration and were not part of the phylogenetic analysis.

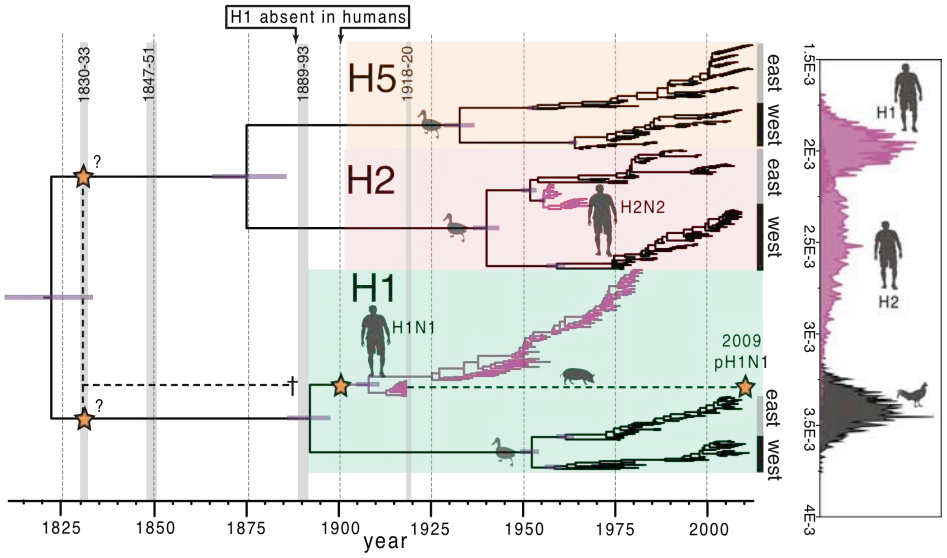
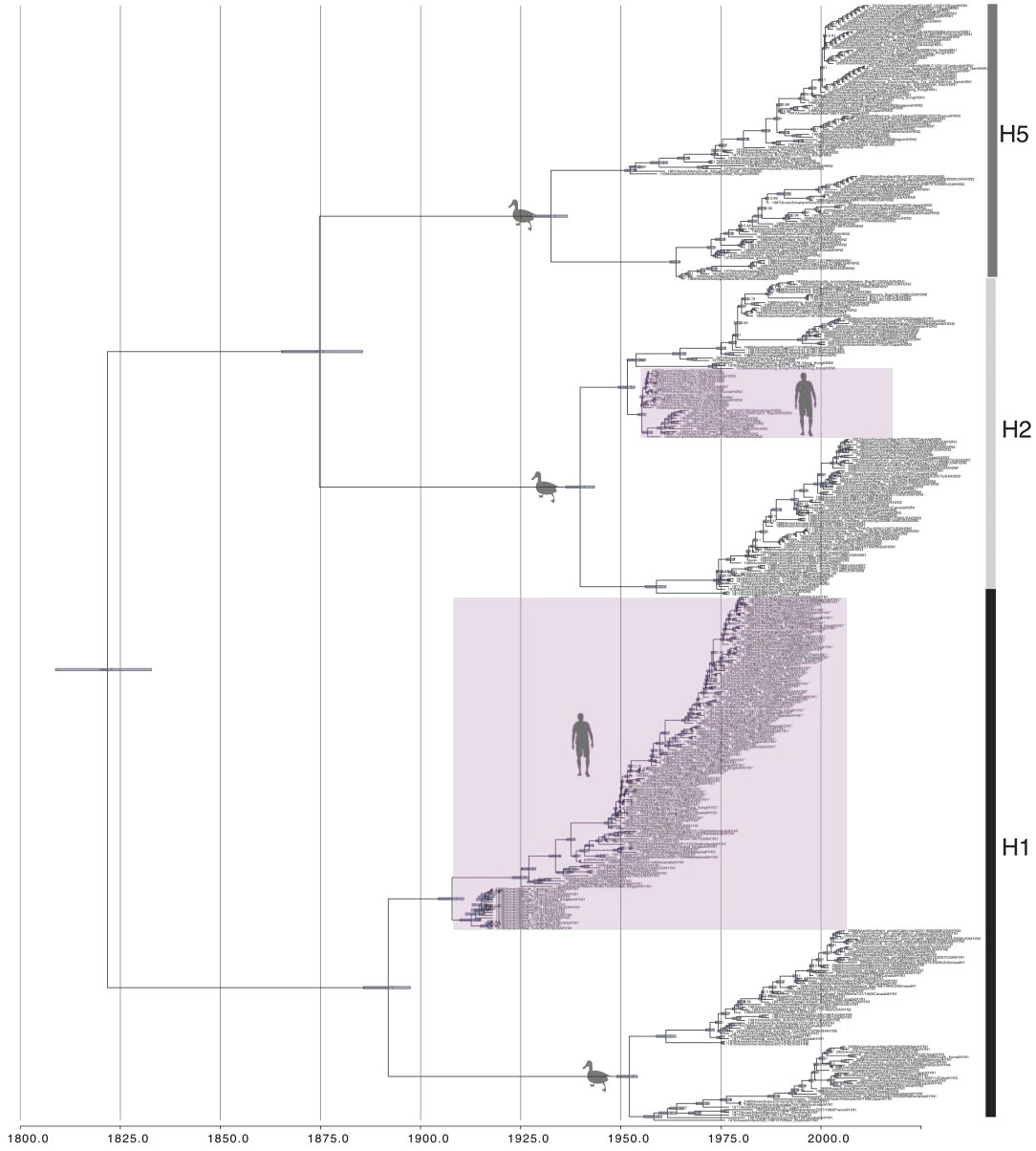


Fig. S14. H1, H2, H5 MCC tree. Branch lengths are in years. Posterior probabilities of each node and 95% CIs on node dates are shown. Human H1 sequences sampled after 1977 (tip-date corrected by -27 years) are marked with an asterisk. (**A**) shared rate for avian H1, H2, and H5. (**B**) independent rates for avian H1, H2, and H5.

A



B

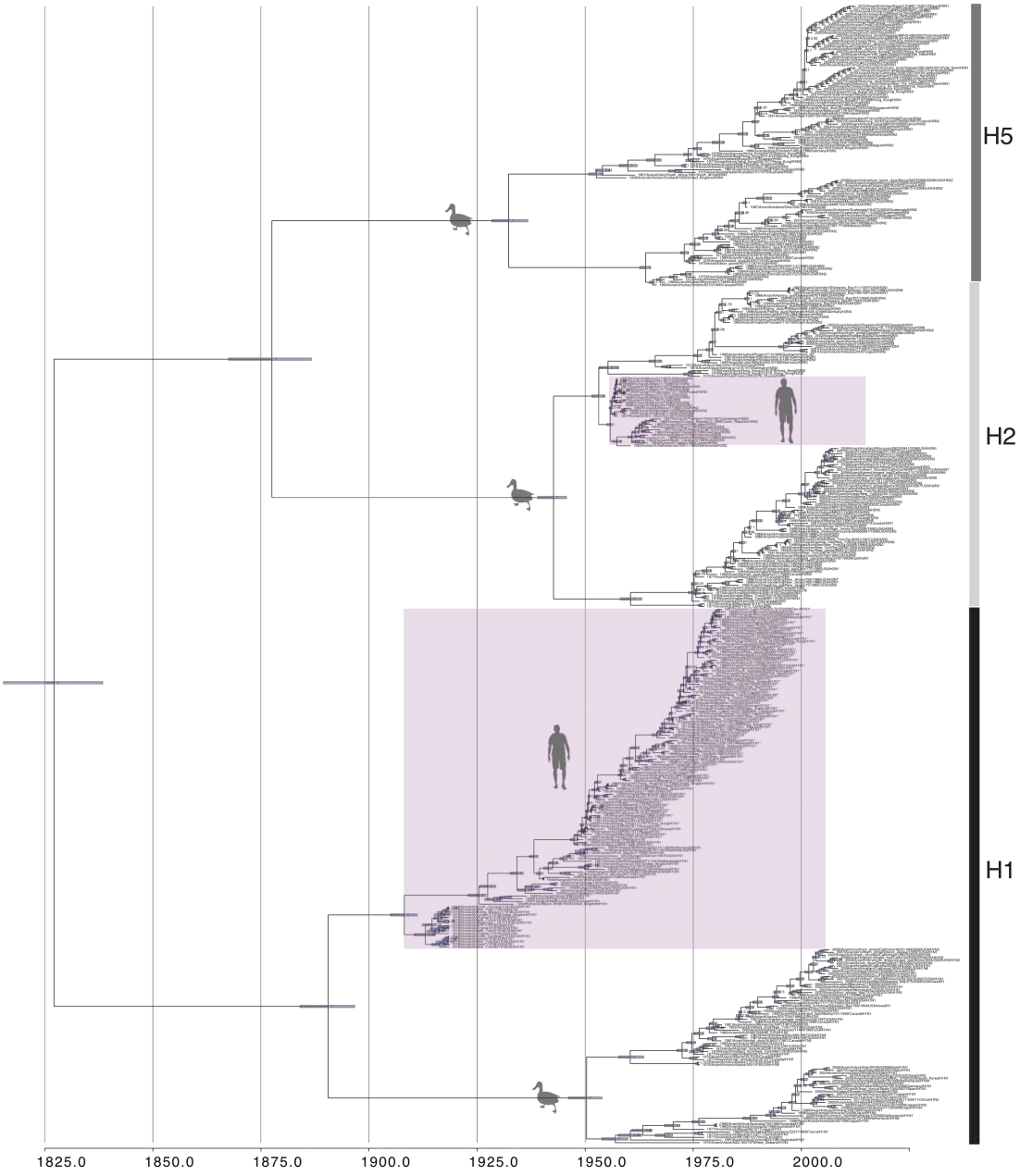


Fig. S15. Fatalities by birth year for H5N1 and H7N9. (**A**) Percentage of fatalities accounted for by 10-year-wide birth year cohorts, for H5N1 and H7N9. Birth year is plotted on the x-axis and the bins are centered on the midpoint of each cohort. Almost all patients with birth years prior to 1968 are expected to have been initially exposed, as children, to a group 1 HA (either H1N1 or H2N2) (blue shading). Most of those born in 1968 and later were exposed first to a group 2 HA (H3N2) (orange shading). (**B**) The same data, but corrected for the percentage of the total population contained in each 10-year birth year cohort in China in 2013 or Indonesia in 2006. See Supplementary text for additional details.